



# Recherche de nouvelles protéines humaines se liant à l'ADN méthylé

Michaël Joulie

## ► To cite this version:

Michaël Joulie. Recherche de nouvelles protéines humaines se liant à l'ADN méthylé. Sciences agricoles. Université Paris Sud - Paris XI, 2011. Français. <NNT : 2011PA112157>. <tel-00660690>

**HAL Id: tel-00660690**

**<https://tel.archives-ouvertes.fr/tel-00660690>**

Submitted on 17 Jan 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THESE

En vue de l'obtention du grade de

## DOCTEUR DE L'UNIVERSITE PARIS XI ECOLE DOCTORALE GENES, GENOMES, CELLULES

Discipline : Science du vivant

Spécialité : Biochimie, Biologie Moléculaire et cellulaire

## RECHERCHE DE NOUVELLES PROTEINES HUMAINES SE LIANT A L'ADN METHYLE

Par Michael JOULIE

Soutenue le

**26 septembre 2011**

### Jury

#### Président

Pr Pierre CAPY

CNRS – Université Paris XI, Gif sur Yvette

#### Rapporteurs

Dr Christian BRONNER

INSERM – Université de Strasbourg, Illkirch

Dr Robert DANTE

CNRS – Université Lyon I, Lyon

#### Examineur

Pr Jonathan WEITZMAN

CNRS – Université Paris VII, Paris

#### Directeur de thèse

Dr Pierre-Antoine DEFOSSEZ

CNRS – Université Paris VII, Paris



## Remerciements

Je remercie les membres du jury d'avoir accepté d'évaluer mon travail, Christian Bronner et Robert Dante, mes rapporteurs, Pierre Capy, le président du jury et Jonathan Weitzman, mon examinateur.

Je remercie particulièrement Pierre-Antoine Defossez pour m'avoir accueilli au sein de son laboratoire. Pierre-Antoine, merci pour la disponibilité dont tu as fait preuve et également de ta volonté à nous faire découvrir les différentes facettes du métier de chercheur.

Merci aux membres passés de l'équipe, Daisuke, Nobu et Véro qui m'ont vu arriver, Moredreck pour sa bonne humeur communicative, Sophie pour ton dynamisme. Et bien sur, merci aux membres présents, Tetris, Benoit et Alex pour la vie que vous apportez au labo.

J'aimerais, également, remercier l'EDC, unité pleine de vie de Johny et Valérie Mezger. Merci à eux et à tous les membres de l'unité, Damien, Mel, Flo, Emma, Marta, Guigui, Claire, Julien et les autres qui participent à l'ambiance conviviale du labo.

Un grand merci aux amis pour tous ces moments passés mais aussi futurs, pour leur soutien et leur amitié. Flo, Had, Pau, Mag, Lolo, Manue, Guille, Sandro...qui me suivent depuis le magistère. Do, c'est un plaisir de t'avoir retrouvé et c'est du bonheur de plus te lâcher. Et merci à Pythagore de nous avoir éclairé un soir d'été.

Pour finir, je remercie, du fond du cœur mes parents pour tout ce que vous m'avez apporté. Merci à ma famille, Dejan, Christine, Josette d'avoir toujours été là et Sylvianne...



# Table des Matières

<b>PARTIE I : INTRODUCTION .....</b>	<b>6</b>
<b>Chapitre 1 : l'épigénétique .....</b>	<b>6</b>
1.1 Evolution du concept d'épigénétique .....	6
1.2 L'ADN support de l'information génétique .....	10
1.3 Les mécanismes impliqués.....	10
1.3.1 La méthylation de l'ADN.....	10
1.3.2 La dynamique des histones.....	12
1.3.3 Les ARNs non codants .....	16
1.3.4 Interactions entre les différentes composantes épigénétiques .....	18
<b>Chapitre 2 : les acteurs moléculaires de la méthylation de l'ADN .....</b>	<b>22</b>
2.1 Mise en place et maintien de la méthylation de l'ADN .....	22
2.1.1 Famille des ADN-méthyltransférases .....	22
2.1.2 DNMT1.....	22
2.1.3 DNMT2.....	27
2.1.4 DNMT3A et DNMT3B .....	28
2.1.5 DNMT3L .....	29
2.2 Interprétation de la méthylation de l'ADN .....	32
2.2.1 Famille à domaine MBD.....	32
2.2.2 Famille à domaine SRA.....	38
2.2.3 Famille de protéines à doigts de Zinc.....	42
<b>Chapitre 3 : rôle de la méthylation de l'ADN .....</b>	<b>46</b>
3.1 Distribution de la méthylation de l'ADN .....	46
3.2 La méthylation de l'ADN au cours du développement.....	48
3.2.1 Reprogrammation de la méthylation de l'ADN des embryons pré-implantatoires.....	48
3.2.2 Reprogrammation de la méthylation de l'ADN des PGCs.....	49
3.3 Contrôle de l'expression de certains gènes tissu-spécifiques.....	50
3.4 Contrôle de l'expression des éléments répétés.....	51
3.4.1 Maintien de l'intégrité du génome.....	51
3.4.2 Rôle dans l'établissement des structures hétérochromatiques.....	52
3.5 Rôle de la méthylation de l'ADN dans le corps des gènes.....	54
3.6 L'empreinte parentale.....	54
3.7 Inactivation du chromosome X.....	58
3.8 Méthylation de l'ADN et maladies humaines.....	58
3.8.1 Défauts de méthylation de l'ADN .....	59
3.8.2 Défauts des acteurs de la méthylation de l'ADN.....	60
3.8.3 Epigénétique et cancer .....	62
3.9 Hydroxyméthylation .....	64
<b>Chapitre 4 : rôle des protéines liant l'ADN méthylé .....</b>	<b>67</b>
4.1 Un dogme à réévaluer.....	67
4.2 Les protéines liant l'ADN méthylé ne sont pas essentielles.....	67
4.2.1 Redondance des MBPs.....	67
4.2.2 Existence d'autres MBPs .....	68
4.3 Cibles des MBPs .....	68
4.4 MBPs : activateurs transcriptionnels ?.....	70

<b>PARTIE II : RECHERCHE DE NOUVELLES PROTEINES HUMAINES SE LIANT A L'ADN METHYLE .....</b>	<b>73</b>
<b>Chapitre 1 : les différentes approches abordées.....</b>	<b>73</b>
1.1 Approche génétique .....	73
1.2 Approche basée sur la littérature .....	74
<b>Chapitre 2 : approche génétique, le phage display .....</b>	<b>76</b>
2.1 Introduction.....	76
2.2 Principe du crible par phage display .....	78
2.3 Développement du crible par phage display .....	79
2.3.1 Choix de l'espèce de bactériophage et de la banque d'ADNc .....	79
2.3.2 Détermination des conditions expérimentales.....	80
2.3.3 Mise au point du crible par phage display.....	84
2.4 Résultats des cribles par phage display.....	86
2.5 Discussion.....	96
2.6 Matériels et méthodes.....	97
<b>Chapitre 3 : approche basée sur la littérature.....</b>	<b>102</b>
3.1 Introduction.....	102
3.2 Résultats.....	104
3.2.1 Les réactifs générés.....	104
3.2.2 Résultats des expériences de colocalisation.....	104
3.2.3 Résultats des expériences de retard sur gel.....	108
3.3 Discussion.....	112
3.3.1 Limites à l'interprétations des résultats.....	112
3.3.2 Interprétations des résultats de colocalisation de BAZ2A.....	113
3.3.3 Perspectives.....	114
3.4 Matériels et méthodes.....	115
<b>Partie III : Régulation des éléments répétés par ZBTB4.....</b>	<b>122</b>
<b>Chapitre 1 Introduction.....</b>	<b>122</b>
<b>Chapitre 2 : résultats.....</b>	<b>124</b>
2.1 Analyse d'expression des éléments répétés .....	124
2.2 Localisation de ZBTB4 <i>in vivo</i> .....	124
2.3 Discussion.....	130
2.4 Matériels et méthodes.....	131
<b>Partie IV : conclusion générale .....</b>	<b>138</b>
<b>Annexe 1.....</b>	<b>142</b>
<b>Annexe 2.....</b>	<b>144</b>
<b>Annexe 3.....</b>	<b>146</b>
<b>Bibliographie .....</b>	<b>158</b>



## Abbreviations

<b>ADN</b>	Acide désoxyribonucléique
<b>mCpG</b>	5-methylcytosine dans un contexte de dinucléotide CG
<b>5hmC</b>	5-hydroxymethylcytosine
<b>5mC</b>	5-methylcytosine
<b>CGI</b>	CpG island
<b>ChIP</b>	chromatine immunoprécipitation
<b>DNMT</b>	ADN méthyltransférase
<b>eGFP</b>	enhanced green fluorescent protein
<b>EMSA</b>	electrophoretic mobility shift assay
<b>cellules ES</b>	cellules souches embryonnaires
<b>GST</b>	glutathione S-transferase
<b>HCPs</b>	high-CpG promoter
<b>HDAC</b>	histone déacetylase
<b>ICP</b>	intermediate-CpG promoter
<b>IgG</b>	immunoglobuline G
<b>kb</b>	1000 (kilo) base-pairs
<b>LCP</b>	Low CpG-promoter
<b>MBD</b>	methyl-CpG binding domain
<b>MBP</b>	methylated-DNA binding protein
<b>MEF</b>	mouse embryonic fibroblast
<b>ARNm</b>	ARN messenger
<b>PCR</b>	polymerase chain reaction
<b>PGCs</b>	primordial germ cells
<b>PCRq</b>	PCR quantitative
<b>ARN</b>	acide ribonucléique
<b>siRNA</b>	small interfering RNA
<b>SRA</b>	SET and RING-associated

**Table 1. Evolution de la biologie au cours du temps**

Biologie		Ere de la génétique	
<b>1800</b> Sens moderne "biologie" (Burdach) <b>1809</b> Concept d'évolution par Lamarck <b>1859</b> <i>L'origine des espèces</i> par Darwin <b>1866</b> Les lois de Mendel		<b>1902</b> Théorie chromosomique de l'hérédité (Sutton et Boveri) <b>1905</b> Mot génétique (Bateson) <b>1911</b> Démonstration théorie chromosomique de l'hérédité (Morgan) <b>1924</b> Les chromosomes sont constitués d'acides nucléiques (Feulgen) <b>1928</b> Transformation génétique (Griffith) <b>1941</b> Un gène, une enzyme (Beadle et Tatum) <b>1944</b> ADN, support de l'hérédité? (Avery, MacLeod et McCarty) <b>1949</b> Equation de Chargaff: complémentarité des bases <b>1952</b> Théorie des protéines porteuses de l'information génétique invalidée (Hershey et Chase)	

Ere de l'ADN		Ere génomique	
<b>1953</b> Structure de l'ADN (Watson, Crick et Franklin) <b>1958</b> Réplication semi-conservative (Meselson et Stahl) <b>1959</b> un gène, un ARNm, une protéine (Jacob et Monod) <b>1965</b> Enzyme de restriction et méthylation de l'ADN (Arber) <b>1965</b> Décryptage du code génétique (Holley, Khorana et Nirenberg)		<b>1972</b> Premier gène séquencé (Fiers) <b>1977</b> Séquençage de l'ADN (Sanger, Gilbert, Maxam) <b>1996</b> Clonage du 1 <sup>er</sup> mammifère <b>2001</b> Séquençage génome humain	

Ere épigénomique	
<b>2006</b> iPS (Yamanaka et Takahashi) <b>2007</b> Etat chromatinien des cellules pluripotentes (Mikkelsen) <b>2009</b> Methylome humain (Lister) <b>2011</b> Analyse dynamique états chromatinien dans 9 types cellulaire humains (Ernst)	

# PARTIE I : INTRODUCTION

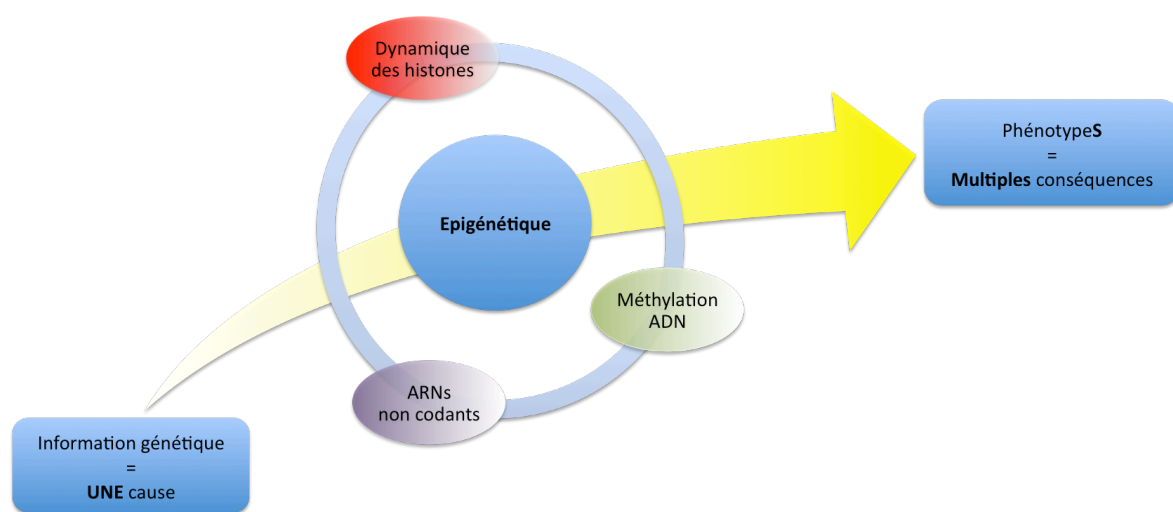
## Chapitre 1 : l'épigénétique

### 1.1 Evolution du concept d'épigénétique

Le concept d'épigénétique émerge à une époque où la génétique réalise une avancée importante grâce aux travaux de Thomas Hunt Morgan, au début du XX<sup>ème</sup> siècle, sur la mutation du gène *white* (Morgan et al. 1910 ; Table 1). Ses recherches permettent de comprendre le rôle joué par les chromosomes dans l'hérédité, alors même que l'acide désoxyribonucléique (ADN) n'est pas caractérisé comme étant le support de l'information génétique. Dès les années 1940, Morgan s'interroge sur le fait que toutes les cellules d'un organisme ne sont pas identiques alors qu'elles partagent les mêmes gènes déterminant les caractères.

Alors que la génétique cherche à comprendre comment sont transmis les caractères, un embryologiste, Conrad Waddington, s'intéresse aux phénomènes développementaux par lesquels les gènes et l'environnement peuvent engendrer un phénotype. Conrad Waddington rassemble alors, en 1942, les théories de la génétique et du développement sous le terme épigénétique (Waddington 1942). Dans sa définition, la notion d'hérédité est cependant absente.

L'importance de l'épigénétique va s'accroître à une période où l'on chercha à comprendre l'activité des gènes. En effet, partant du constat qu'une cellule œuf est capable d'engendrer tous les types cellulaires d'un organisme, comment expliquer qu'une cellule neuronale soit différente d'une cellule musculaire alors même qu'elles partagent la même information génétique? Cela signifie qu'un certain nombre de gènes, spécifiques d'un lignage cellulaire, sont activés tandis que d'autres sont réprimés. Comment ces gènes sont-ils alors régulés? Une première partie de la réponse fût apportée par les travaux de François Jacob et de Jacques Monod sur la régulation de l'opéron lactose (Jacob et Monod 1961). De ce modèle, ils en déduisent l'existence deux classes de gènes : les gènes structuraux et les gènes régulateurs qui agissent de façon séquence spécifique.



### Figure 1. L'épigénétique

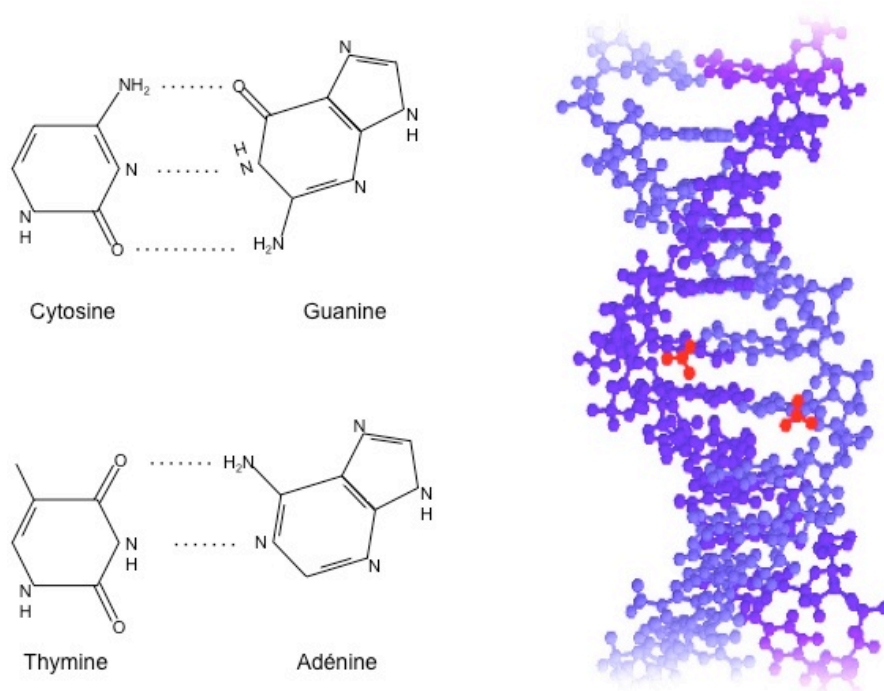
Les mécanismes épigénétiques sont principalement les histones et leur dynamique au sein de la cellule, les ARNs non codants et la méthylation de l'ADN. Ces mécanismes permettent d'interpréter de différentes manières une seule et unique information. L'épigénétique apporte un degré supplémentaire quant à la régulation de l'activité de notre génome participant ainsi à la compréhension de la complexité de l'espèce humaine.

Cependant, cette vision se heurte à quelques phénomènes non explicables. L'inactivation du chromosome X chez les mammifères en est un exemple. En 1959, Susumu Ohno remarque que le corpuscule de Barr est en réalité un des chromosomes X condensé (Ohno et al. 1959). Ce n'est que deux ans plus tard, en 1961, que Mary F. Lyon propose l'hypothèse qu'un des deux chromosomes X chez les mammifères femelles est inactivé de manière aléatoire (Lyon 1961). Cette théorie est contradictoire avec la vision de régulation des gènes de Jacob et Monod. Comment peut-on expliquer que des facteurs de transcription inactivent un seul des deux chromosomes X sachant que les séquences cibles sont présentes à la fois sur le chromosome paternel et maternel.

Par ailleurs, un autre phénomène montre les limites du système gènes structuraux/gènes régulateurs. En 1984, la mise au point d'expériences de transplantations nucléaires chez la souris (McGrath et Solter 1983 ; Surani et al. 1984) a permis de mettre en évidence que des embryons de souris androgénètes (deux génomes paternels) ou gynogénètes (deux génomes maternels) étaient incapables de se développer. Solter et McGrath conclurent que la contribution du génome paternel et celle du génome maternel ne sont pas équivalentes au cours de l'embryogenèse chez les mammifères, posant ainsi les bases de ce que l'on appelle aujourd'hui l'empreinte parentale.

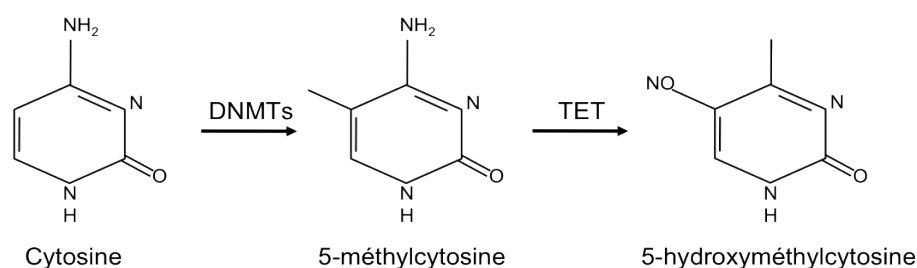
Bien que ne faisant pas l'objet d'un consensus total (Bird et al. 2007 ; Berger et al. 2009), ces phénomènes inexpliqués ont permis d'établir la définition moderne du terme épigénétique comme étant l'étude des phénomènes qui modifient l'activité du génome, de façon héritable au cours de la mitose et/ou de la méiose, et qui n'affectent pas la séquence d'ADN.





**Figure 2. Structure de l'acide désoxyribonucléique.**

La complémentarité de l'ADN est rendue possible par les liaisons hydrogènes entre les bases A-T et G-C. La molécule d'ADN est une double hélice antiparallèle. La cytosine peut être modifiée par l'ajout d'un groupement méthyle (en rouge). La molécule est méthylée symétriquement sur un CpG. Les méthyles occupent une position accessible dans le grand sillon et peuvent modifier les contacts entre les protéines et l'ADN.



**Figure 3. Modifications de la cytosine chez les mammifères.**

La cytosine peut être modifiée par l'ajout d'un groupement méthyle en position 5 grâce à des enzymes spécialisées appelées ADN méthyltransférases (DNMTs). La 5-méthylcytosine peut alors être hydroxylée en 5-hydroxyméthylcytosine par l'enzyme TET.

## 1.2 L'ADN support de l'information génétique

En 1953, Watson et Crick, avec l'aide des travaux de Rosalind Franklin, parviennent à résoudre la structure de l'ADN comme étant une double hélice formée de deux brins complémentaires antiparallèles (Watson et Crick, 1953). L'ADN est un polymère de nucléotides constitué d'un groupe phosphate lié à un désoxyribose lui même rattaché à une base azotée. Il existe deux types de bases azotées (figure 2) :

- Les pyrimidines comprenant la thymine (T) et la cytosine (C) et l'uracile (U) pour l'ARN.
- Les purines comprenant l'adénine (A) et la guanine (G).

La complémentarité de l'ADN est rendue possible par les interactions entre les bases A-T et G-C. Ces observations permettent alors d'émettre l'hypothèse que l'ADN est le support de l'information génétique et que l'hérédité peut être expliquée par un mécanisme de séparation des brins d'ADN suivi d'une duplication. Matthew Meselson et Franklin Stahl démontrent cette théorie en établissant le modèle de répllication semi-conservative en 1958 (Meselson et Stahl, 1958). Dès lors, les scientifiques se penchent sur le déchiffrement du code génétique : comment un brin d'ADN peut-il être traduit en protéines ? Et dès les années 1960, Robert Holley (Holley et al. 1965a ; Holley et al. 1965b), Har Gobind Khorana (Kornberg et al. 1964) et Marshall Nirenberg (Matthaei et al. 1962) résolvent ce code.

## 1.3 Les mécanismes impliqués

Les phénomènes concernés par l'épigénétique dépendent de la définition utilisée. Je prendrai ici en compte la dynamique des histones, l'intervention des ARNs non codants et la méthylation de l'ADN (figure 1). Cette dernière est certainement le plus clairement établi, car son hérédité ne fait aucun doute.

### 1.3.1 La méthylation de l'ADN

Outre les quatre bases A,T,G et C qui composent la molécule d'ADN, on peut noter l'existence de modifications chimiques sur ces bases. Parmi ces modifications, on retrouve, chez les mammifères, l'ajout d'un groupement méthyle en position 5 d'une cytosine grâce à des enzymes spécialisées

appelées ADN méthyltransférases (DNMTs ; figure 3). La cible principale de la méthylation chez les mammifères est la cytosine dans un contexte de dinucléotide CpG. Cependant, il a été récemment montré, dans des cellules souches embryonnaires, que la méthylation peut également avoir lieu sur des trinucléotides CHH et CHG (où H = A,C ou T) (Lister et al. 2009) et dans un contexte non-CpG, principalement CpA et CpT (Ramsahoye et al. 2000 ; Laurent et al. 2010). Les dinucléotides CpGs ne sont pas distribués de façon homogène au sein du génome mais ont tendance à être concentrés dans des régions appelées îlots CpG (CGIs pour CpG islands). Ces régions sont définies par un contenu de G+C d'au moins 50% sur une longueur de plus de 200 paire de bases (pb) et par un ratio de CpG observé par rapport à la fréquence attendue d'au moins 0,6. La fréquence globale des dinucléotides CpGs est ainsi très faible ( $\approx 1\%$ ) tandis que, chez l'Homme, environ 60% des promoteurs de gènes humains possèdent un CGI. Bien que celui-ci soit habituellement non méthylé dans les cellules normales. On note tout de même qu'environ 6% des CGIs deviennent méthylés de manière tissu-spécifique au cours du développement ou de la différenciation (Straussman et al. 2009 ; Illingworth et al. 2008 ; Illingworth et al. 2010 ; Borgel et al. 2010).

Les observations effectuées dans les cellules cancéreuses permettent d'établir une bonne corrélation entre la méthylation des promoteurs et la répression transcriptionnelle. En effet, les gènes des cellules tumorales dont les promoteurs à CGIs sont hyperméthylés, affichent un état transcriptionnellement inactif. Ce phénotype est réversible puisque la suppression de la méthylation de l'ADN, par l'utilisation de drogue inhibant les DNMTs, permet la réexpression des gènes touchés (Esteller 2008).

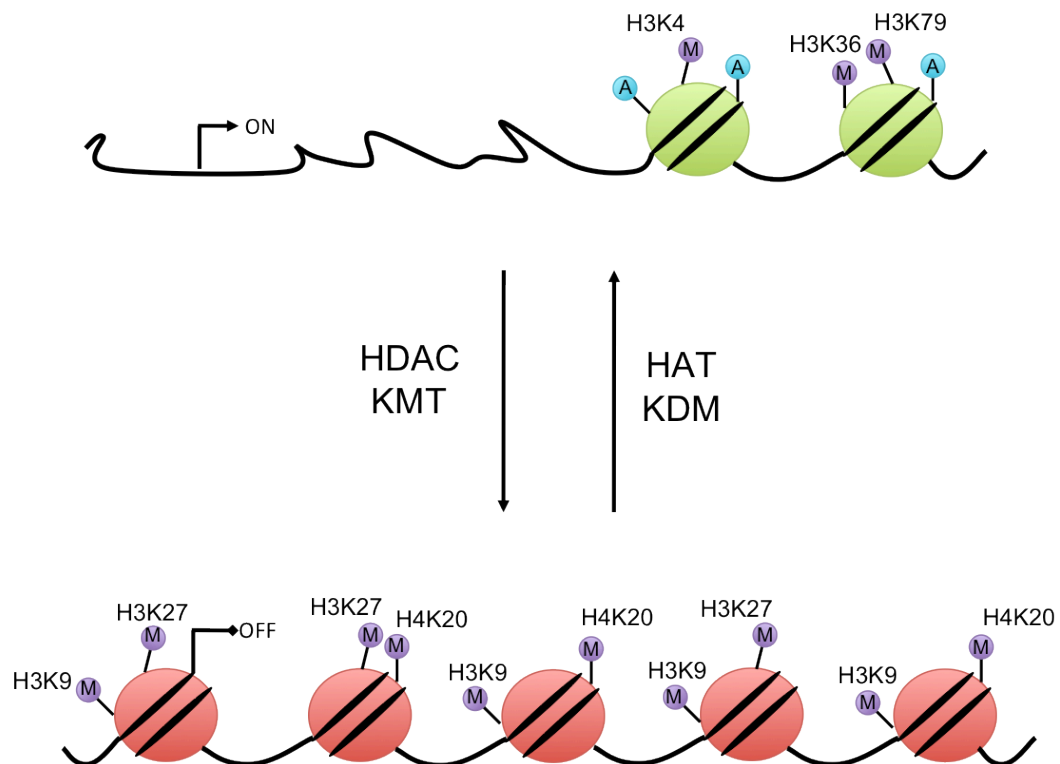
Plus récemment, des travaux ont établi que cette corrélation n'est pas uniquement restreinte aux cellules tumorales puisqu'elle concerne également les gènes de la lignée germinale. Ces derniers possèdent des promoteurs à CGIs qui sont non-méthylés dans les cellules germinales mais méthylés dans les cellules somatiques (Magdinier et al. 2002). La répression, dans ce cas précis, nécessite donc la méthylation de l'ADN (Maatouk et al.

2006 ; Velasco et al. 2010). On peut également citer l'exemple des gènes du trophoctoderme qui sont réprimés dans l'embryon par une hyperméthylation des CGIs (Oda et al. 2006).

### 1.3.2 La dynamique des histones

Les histones sont des protéines basiques de faible poids moléculaire qui permettent la condensation de l'ADN. La structure ainsi formée est appelée nucléosome. Chaque nucléosome est composé d'un octamère d'histone comprenant deux copies de chaque histone H2A, H2B, H3 et H4. Environ 147 pb d'ADN sont enroulées autour de chaque nucléosome (Luger et al. 1997). Enfin, l'histone H1 sert de liaison afin de sceller le nucléosome. Les histones sont constituées de deux parties structurellement bien distinctes, le corps des histones et la queue des histones. Cette dernière partie est sujette à de nombreuses modifications post-traductionnelles (figure 4) : phosphorylation, acétylation, méthylation, SUMOylation, ubiquitination, ADP-ribosylation, déamination, clivage de la queue et isomérisation de la proline (Bannister et Kouzarides 2011). Ces modifications sont mises en place par des protéines spécialisées : histones acétyltransférases (HATs) (Sternier et Berger 2000), histones méthyltransférases (HMTs) (Zhang et Reinberg 2001), ou encore kinases (Nowak et Corces 2004) et possèdent la propriété d'être dynamiques et donc, réversibles grâce à des enzymes comme les histones déacétylases (HDAC). Des histones déméthylases (HDMs) ont également été mises en évidence récemment pour la lysine (Shi et al. 2004 ; Whetstine et al. 2006) ainsi que pour l'arginine (Chang et al. 2007).

Contrairement à la méthylation de l'ADN, les mécanismes de transmission des modifications des histones au cours des divisions cellulaires ne sont pas encore totalement clairs. Cependant, il semble que le complexe homotrimérique PCNA est capable de recruter des HDACs, des lysines méthyltransférases (KMTs), des enzymes de remodelage de la chromatine ainsi que CAF1 au niveau de la fourche de réplication (Probst et al. 2009). Par ailleurs, le complexe DNMT1-UHRF1 recrute également des HDACs et des KMTs (Fuks et al. 2000 ; Rountree et al. 2000 ; Unoki et al. 2004 ; Estève et al. 2006). Ces données permettent une meilleure compréhension de la



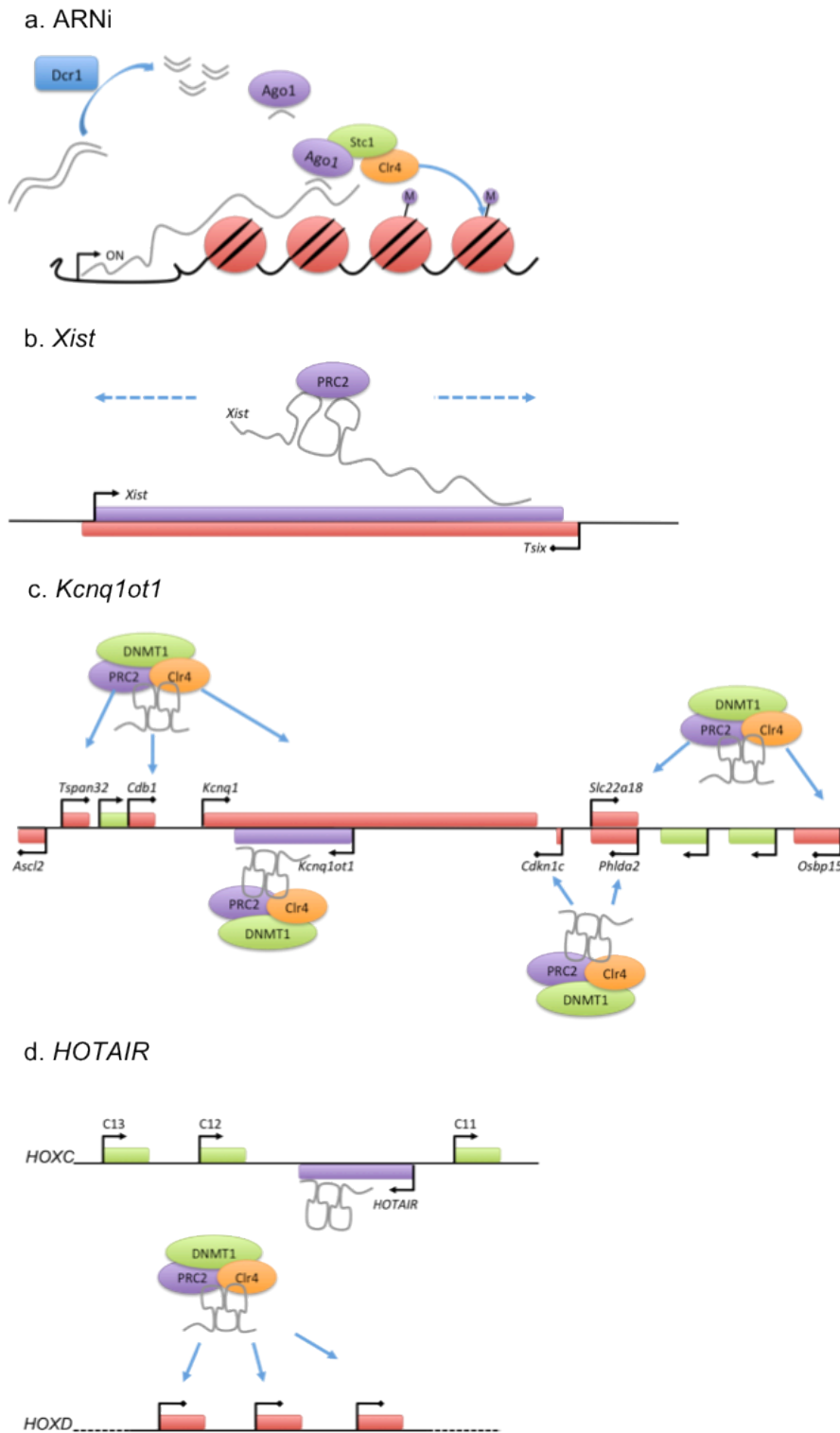
**Figure 4. La dynamique des histones.**

Les queues des histones peuvent subir des modifications post-traductionnelles. Ces modifications peuvent être associées à un état transcriptionnellement actif avec la méthylation de H3K4, H3K36 et H3K79 et avec l'acétylation des histones. En revanche, la méthylation de H3K9, H3K27, H4K20 est assimilée à un état répressif. Ces marques sont dynamiques et régulées par des enzymes tels que les histones déacétylases (HDAC), les lysines méthyltransférases (KMT), les histones acétyltransférases (HAT) ou encore les lysines déméthylase (KDM).

dynamique et de la stabilité des modifications des histones au cours du cycle cellulaire.

Les modifications des histones ont un rôle important à jouer, d'une part dans la régulation de l'activité des gènes et d'autre part, dans des processus du métabolisme de l'ADN tels que la réparation, la réplication et la recombinaison. Deux mécanismes distincts interviennent dans la fonction des modifications des histones. Le premier, de manière directe perturbe la structure de la chromatine. Par exemple, l'acétylation et la phosphorylation ont pour effet de diminuer la charge positive des histones et perturbent ainsi les interactions ADN-histones, résultant alors en une décondensation locale de la chromatine et donc en un état permissif pour la transcription. Le deuxième se fait quant à lui de manière indirecte via le recrutement de cofacteurs. C'est notamment le cas de HP1 capable de lier, d'une part, la triméthylation de la lysine 9 (H3K9me3) par son chromodomaine (Bannister et al. 2001 ; Lachner et al. 2001 ; Jacobs et Khorasanizadeh 2002) et, d'autre part, d'interagir avec Suv39h1 (Fujita et al. 2003), les ADN méthyltransférases (Fuks et al. 2003), les MBPs (Fujita et al. 2003 ; Agarwal et al. 2007) ainsi que les HDACs (Zhang et al. 2002). Ces observations suggèrent l'existence de mécanismes de mise en place et de maintien des marques répressives de la chromatine (Nielsen et al. 2001).

Historiquement, la chromatine a été caractérisée, selon sa structure, en euchromatine et en hétérochromatine. L'euchromatine correspond à une structure de l'ADN peu compacte en interphase, dite en collier de perle, regroupant des gènes transcriptionnellement actifs. Tandis que l'hétérochromatine présente un niveau de compaction plus important en fibre d'ADN de 30nm. L'hétérochromatine peut être divisée en hétérochromatine constitutive, constituée de peu de gènes et principalement de séquences répétées retrouvées au niveau des centromères et des télomères, et en hétérochromatine facultative dont le degré de compaction diffère d'un type cellulaire à un autre. Les marques représentatives de l'euchromatine sont l'acétylation des histones, H3K4me3, H3K36me3 et H3K79me3, alors que l'hétérochromatine est, quant à elle, caractérisée par un faible taux d'acétylation ainsi que H3K9me, H3K27me et H4K20me.



**Figure 5. Les ARNs non codants.**

**A :** Mécanisme d'interférence à l'ARN nécessaire à l'initiation de la formation d'hétérochromatine constitutive chez la levure *Schizosaccharomyces pombe*. La transcription mène à la formation d'un ARN double brin qui est dégradé en siARN par l'enzyme Dcr1. La protéine Ago1 lie ces petits ARNs et cible les sites de transcription. L'histone méthyltransférase Clr4 est alors recruté via Stc1 et permet la formation d'un état chromatinien répressif. **B :** L'inactivation du chromosome X nécessite l'accumulation d'un long ARN non codant *Xist* et le recrutement du complexe Polycomb PRC2 afin d'induire la répression transcriptionnelle. **C :** De long ARNs non codants, tel que *Kcnq1ot1*, régulent *en cis* l'expression des gènes soumis à l'empreinte via le recrutement de l'histone méthyltransférase G9a (non présenté), du complexe PRC2 et de l'ADN méthyltransférase DNMT1. **D :** L'ARN non codant *HOTAIR* exprimé à partir du locus *HOXC* induit la répression transcriptionnelle *en trans* du locus *HOXD*.

Bien que les modifications individuelles puissent être associées à un niveau d'activité de certaines régions du génome, la signification biologique des combinaisons des différentes marques reste énigmatique. Sachant qu'il existe plus de cent modifications d'histones distinctes, les notions d'euchromatine et d'hétérochromatine d'un point de vue fonctionnel doivent donc être réévaluées. En ce sens, de nouvelles approches cherchent à découvrir de nouveaux états chromatinien et à comprendre la répartition spatiale des marques combinées (Ernst et Kellis 2010). A ce jour, cinquante-et-un états chromatinien différents ont été identifiés à partir de l'enrichissement de combinaisons spécifiques de modifications d'histones. Et ces différents états ont pu être associés à une fonction biologique. De la même manière, la compréhension de la diversité cellulaire au sein d'un organisme passe par la caractérisation systématique des différentes marques chromatinien dans une multitude de type cellulaires (Ernst et al. 2011).

### 1.3.3 Les ARNs non codants

Durant des décennies, l'ARN a été considéré comme simple intermédiaire à la synthèse des protéines en accord avec le dogme central « un gène, un ARN, une protéine ». En 1972, Susumu Ohno introduit la notion d'ADN « poubelle » comme étant les portions du génome sans fonction biologique. Il s'avère désormais que des régions du génome peuvent être transcrites en ARN sans pour autant coder pour des protéines, on parle alors d'ARNs non codants. Les avancées dans ce domaine ont permis de montrer que ces ARNs non codants ont en réalité diverses fonctions nécessaires à la régulation de la chromatine, l'organisation nucléaire mais également à la régulation de la transcription (Beisel et Paro 2011 ; figure 5).

En 1998, Fire et Mello identifient un mécanisme par lequel des petits ARNs fournissent une spécificité de séquence permettant la régulation de l'expression de gènes cibles, c'est la découverte de l'ARN interférent (ARNi) (Fire et al. 1998 ; figure 5a). Chez la levure *Schizosaccharomyces pombe*, ces petits ARNs permettent également la formation d'hétérochromatine via les interactions de la machinerie du ARNi avec des HMTs conduisant alors à la répression transcriptionnelle (Moazed 2009). La formation



d'hétérochromatine dépendante de l'ARNi est également retrouvée chez *Arabidopsis thaliana*, *Caenorhabditis elegans* et *Drosophila melanogaster*. En revanche, cela n'a pas encore été démontré chez les mammifères. Cependant, chez ces derniers, de longs transcrits non codants contribuent à la régulation des gènes soumis à l'empreinte, du développement ou encore de la compensation de dose lors du phénomène d'inactivation du chromosome X. Pour cela, les longs ARNs non codants participent aux changements épigénétiques en recrutant des complexes de remodelage de la chromatine à des loci spécifiques (Tsai et al. 2010 ; Nagano et al. 2008).

Les longs ARNs non codants peuvent agir en *cis* comme l'ARN *Xist* qui est exprimé à partir du chromosome X inactif et qui agit sur l'ensemble du chromosome (Brockdorff et al. 1991 ; figure 5b). *Xist* est un ARN d'environ 19 kb qui s'accumule autour du chromosome X inactif et induit des modifications de la chromatine conduisant à la répression transcriptionnelle du chromosome X entier (voir chapitre 3.7).

D'autres ARNs non codants tels que *Air* et *Kcnq1ot1* interviennent sur une région plus restreinte permettant le contrôle respectivement des gènes soumis à l'empreinte des loci *Igf2r* et *Kcnq1* (figure 5c). L'ARN *Kcnq1ot1* est long de 90 kb et interagit avec l'histone méthyltransférase G9a, le complexe Polycomb PRC2 (Pandey et al. 2008) et l'ADN méthyltransférase DNMT1 (Mohammad et al. 2010) formant alors une structure chromatinienne non permissive pour la transcription. *Igf2r* est exprimé maternellement, l'allèle paternel doit donc être réprimé. Pour cela, l'hypométhylation de l'allèle paternel est nécessaire à la transcription du long ARN non codant *Air* permettant l'inhibition du locus *Igf2r* (Sleutels et al. 2002).

Par ailleurs, les ARNs non codants peuvent également agir en *trans*, c'est le cas des gènes Hox où un ARN non codant, *HOTAIR*, transcrit à partir du locus *HOXC*, va induire en *trans* la répression du gène *HOXD* situé à 40 kb (Rinn et al. 2007 ; figure 5d). *HOTAIR* interagit avec le complexe PRC2 qui est nécessaire à la méthylation de l'histone H3 du locus *HOXD* mais également avec la lysine déméthylase LSD1 et les complexes CoREST et REST (Tsai et al. 2010).

Outre les longs ARNs non codants, de petits ARNs de 25 à 28 nt (piARNs) participent à la régulation des éléments transposables dans les cellules germinales des mammifères. Ces petits ARNs forment un complexe avec des protéines de la famille PIWI afin de cliver les transcrits provenant des éléments transposables et d'induire également une inhibition transcriptionnelle (Aravin et al. 2008). Bien que le lien entre piARNs et méthylation de l'ADN ne soit pas clairement établi, il apparaît de façon certaine que DNMT3A et DNMT3B jouent un rôle dans le processus de méthylation et de répression de ces éléments transposable (voir chapitre 2.1.4).

Enfin, les ARNs non codants, de par leur capacité à reconnaître des séquences spécifiques par complémentarité, sont amenés à réguler différents processus post-transcriptionnels tels que l'épissage (Beltran et al. 2008).

#### 1.3.4 Interactions entre les différentes composantes épigénétiques

La méthylation de l'ADN est liée à la fois mécanistiquement et fonctionnellement aux autres composantes épigénétiques et notamment aux marques d'histones. De façon générale, la méthylation de H3K4, H3K36 et H3K79 est associée à un état transcriptionnellement actif tandis que la méthylation de H3K9, H3K27 et H4K20 est corrélée à un état inactif de la chromatine (Kouzarides 2007). Par ailleurs, la méthylation de l'ADN et le variant d'histone H2A.Z sont mutuellement exclusif (Zilberman et al. 2008).

Afin de comprendre les relations entre méthylation de l'ADN et méthylation des histones, les profils de méthylation de l'ADN de séquences répétées ont été étudiés dans les cellules ES murines délétées pour *Suv39h*, une méthyltransférase de H3K9. Il s'avère que la perte de H3K9me3 provoque une hypométhylation des répétitions péri-centriques tandis que les autres éléments conservent leur niveau de méthylation de l'ADN (Lehnertz et al. 2003). En revanche le niveau de H3K9me3 demeure inchangé dans les cellules ES invalidées pour les DNMTs. Ces résultats suggèrent alors que la méthylation des histones est l'événement initiateur conduisant à la

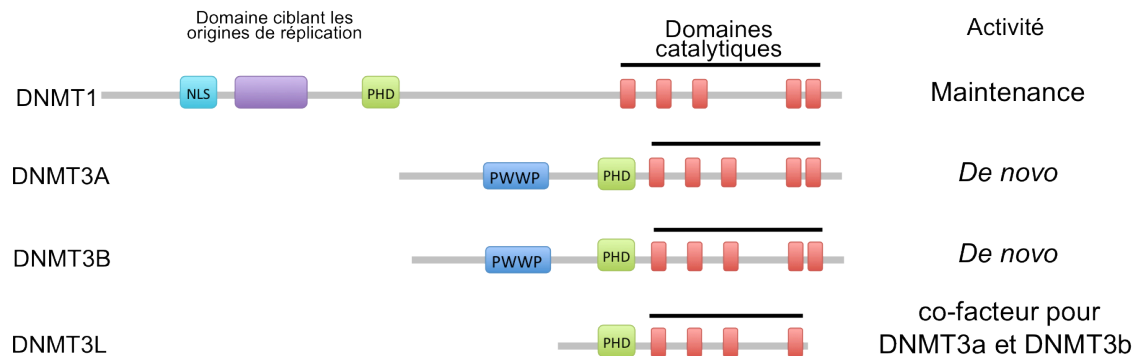
méthylation de l'ADN au niveau des répétitions péricentriques chez la souris.

Il a également été montré que HP1 permet de connecter la méthylation de H3K9 avec la méthylation de l'ADN. HP1 possède un chromodomaine liant H3K9 méthylé et permettant le recrutement des DNMTs à l'hétérochromatine interagit avec les différentes DNMTs (Smallwood et al. 2007).

A l'inverse de la méthylation d'H3K9, H3K4 méthylé semble inhiber la méthylation de l'ADN. DNMT3L, le co-facteur des ADN méthyltransférases *de novo*, lie la queue des histones H3 mais cette interaction est soulevée lorsque H3K4 est méthylé, suggérant ainsi le rôle de la méthylation de H3K4 dans l'établissement de la méthylation de l'ADN (Ooi et al. 2007). Il apparaît d'ailleurs que la méthylation H3K4 doit être effacée par la déméthylase KDM1B afin de permettre la méthylation *de novo* de l'ADN sur les gènes soumis à l'empreinte maternelle (Ciccone et al. 2009).

D'autre part, un lien entre la méthylation de l'ADN et l'acétylation des histones est fortement suggéré par l'anti-corrélation existant entre ces deux marques. Cependant, si la connexion, d'un point de vue fonctionnel sur la régulation de l'expression des gènes est établie, la question sur les mécanismes mis en cause reste en suspens. Les MBPs, et en particulier MeCP2, sont capables d'interagir avec le complexe Sin3A et des HDACs confirmant ainsi la relation étroite entre les deux types de modifications (Nan et al. 1998 ; voir Chapitre 2.2.1). De plus, le traitement de lignées cellulaires cancéreuses par des inhibiteurs de HDAC tel que le Valproate (VPA) ou la Trichostatin A (TSA) provoque une diminution de la méthylation de l'ADN (Milutinovic et al. 2007 ; Ou et al. 2007).





**Figure 6. Structure des différentes protéines méthylant l'ADN.**

DNMT1 permet le maintien de la méthylation tandis que DNMT3a, DNMT3b et DNMT3L ont un rôle dans l'établissement de la méthylation. Les différentes ADN méthyltransférases partagent un domaine PHD et un domaine catalytique. Cependant, le domaine catalytique de DNMT3L est tronqué et non fonctionnel. DNMT1 possède un domaine ciblant les origines de réplication. DNMT3A et DNMT3B possèdent un domaine PWWP.

## Chapitre 2 : les acteurs moléculaires de la méthylation de l'ADN

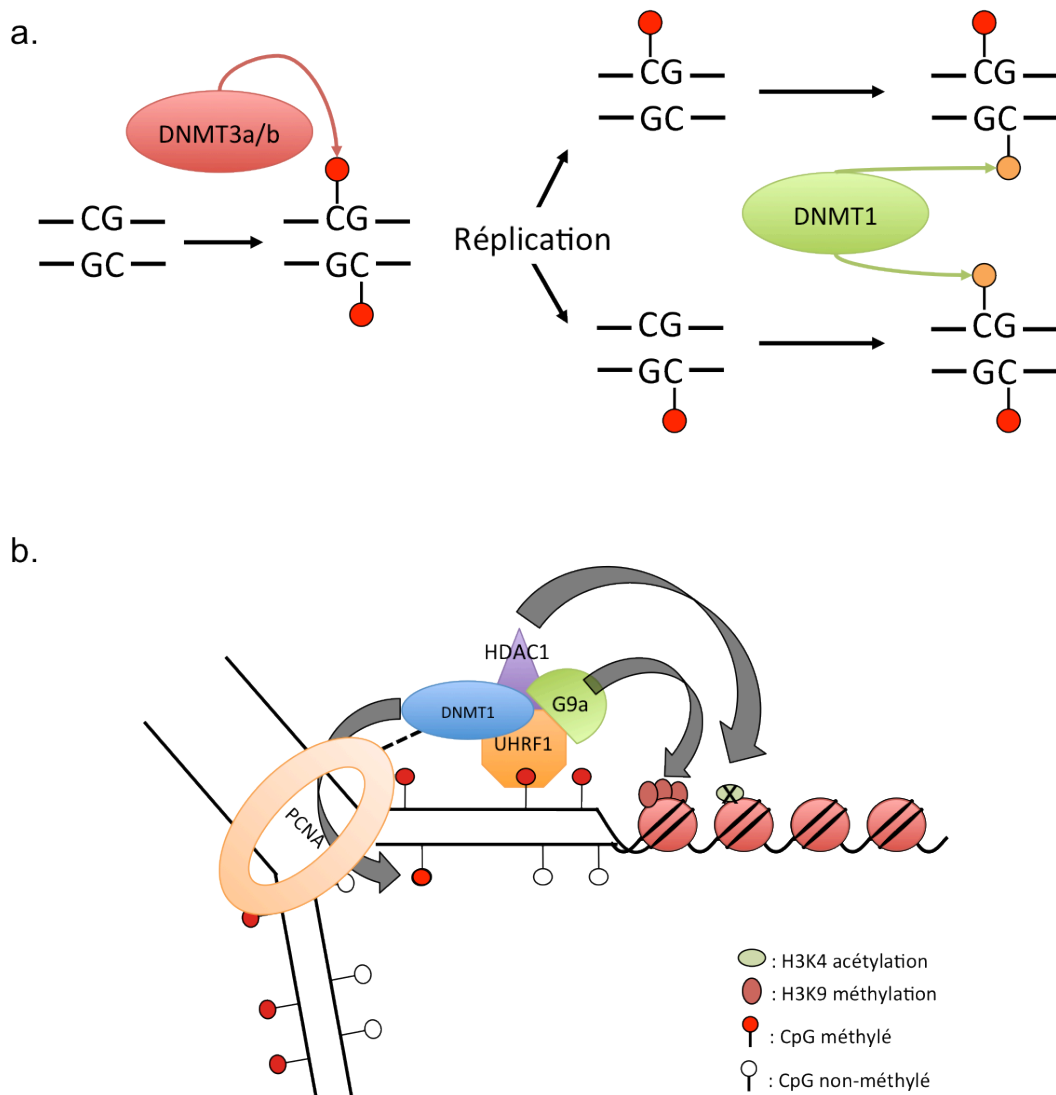
### 2.1 Mise en place et maintien de la méthylation de l'ADN

#### 2.1.1 Famille des ADN-méthyltransférases

##### 2.1.2 DNMT1

En 1988, l'ADNc codant pour l'ADN méthyltransférase 1 (DNMT1) de souris a été cloné et séquencé. La séquence protéique de la région C-terminale présente une homologie avec les DNMTs bactériennes suggérant qu'il s'agit du domaine catalytique (Bestor et al. 1988 ; figure 6). La caractérisation biochimique a été permise par la purification de la protéine entière démontrant son affinité pour l'ADN hémiméthylé *in vitro* (Pradhan et al. 1999). DNMT1 est depuis lors considérée comme le facteur de maintenance de la méthylation de l'ADN (figure 7a). Cette découverte confirme le modèle prédit près de 25 ans plus tôt par Holliday, Pugh et Riggs en 1975 qui estiment qu'il existe deux type d'ADN méthyltransférases : les DNMTs *de novo* et les DNMTs avec une activité de maintenance. *In vivo*, la fonction de maintenance est confirmée par les données d'expression de *Dnmt1* au cours du cycle cellulaire, de localisation de la protéine et de par ses interactions. En effet, l'augmentation de l'expression de *Dnmt1* durant la phase S est corrélée avec la maintenance de la méthylation de l'ADN au cours de la duplication (Roberston et al. 2000). Ensuite, DNMT1 localise au niveau des fourches de réplication (Leonhardt et al. 1992). Enfin, DNMT1 interagît directement avec la machinerie de duplication via PCNA (Chuang et al. 1997) ainsi qu'avec UHRF1, une protéine qui lie l'ADN hémiméthylé (Bostick et al. 2007 ; Sharif et al. 2007 ; figure 7b).

La mutation du gène *Dnmt1* dans les embryons de souris provoque une létalité embryonnaire entre 12,5 – 15,5 jours de gestation (Li et al. 1992). DNMT1 est donc essentielle au développement au cours de l'embryogenèse. Cependant, les cellules souches embryonnaires mutées pour *Dnmt1* sont viables mais si elles présentent toujours un niveau faible de méthylation suggérant l'existence d'au moins une autre DNMT (Lei et al. 1996).



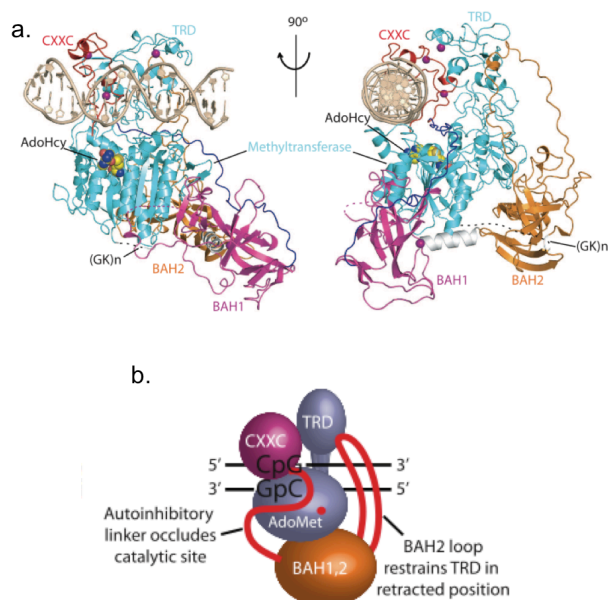
**Figure 7. Activité *de novo* et de maintenance des ADN méthyltransférases.**  
**A :** Les ADN méthyltransférases. DNMT3A et DNMT3B ont été identifiés pour avoir essentiellement une activité *de novo* et permettent la mise en place des profils de méthylation de l'ADN. DNMT1 est l'ADN méthyltransférase de maintenance et rétabli les profils de méthylation après le passage de la fourche de réplication. **B :** Modèle du processus de maintenance des marques épigénétiques. DNMT1, de par son interaction avec UHRF1 et PCNA, est recruté au niveau des sites hémiméthylés au cours de la réplication et maintien la méthylation de l'ADN. Le complexe comprend également HDAC1 et G9a afin de conserver l'état chromatinien. Source : Sharif et al. 2007

De nombreuses études se sont intéressées aux partenaires de DNMT1 pour comprendre sa fonction. A ce jour, 31 protéines ont été identifiées comme jouant un rôle dans le complexe de réplication (PCNA, UHRF1, HP1 $\beta$ , G9a, SUV39H1, HDACs...), dans la réparation de l'ADN (PARP-1), dans le contrôle du cycle cellulaire (pRb/E2F1) ou encore dans la régulation de l'apoptose (p53). Il a également été montré que DNMT1 peut avoir une activité transcriptionnelle répressive (Fuks et al. 2003 ; Viré et al. 2006). Afin de dissocier le rôle de la méthylation du rôle de DNMT1 dans le développement embryonnaire, des expériences de sauvetage du phénotype mutant pour DNMT1 dans les cellules ES murines ont été réalisées (Damelin et Bestor 2007). Lorsqu'une mutation est introduite dans le site catalytique de DNMT1, la protéine mutée ne parvient pas à sauver le phénotype des cellules ES *Dnmt1*<sup>-/-</sup>, indiquant que la méthylation de l'ADN est essentielle et que les fonctions biologiques de DNMT1 passent par son activité méthyltransférase. La génération de souris mutantes pour le domaine catalytique de DNMT1 confirme l'importance de la méthylation de l'ADN (Takebayashi et al. 2007).

L'analyse des structures cristallographiques des DNMTs bactériennes montrent que les DNMTs ont besoin d'au moins trois éléments structuraux pour leur activité catalytique : un domaine de liaison au co-facteur S-Adenosyl-méthionine (SAM) donneur de méthyle, une cavité permettant la liaison de la cytosine cible à méthyler et un domaine de reconnaissance des séquences cibles. C'est l'homologie existante entre les DNMTs bactériennes comme Hha1 et DNMT1 qui ont permis de déduire les mécanismes catalytiques de DNMT1. Afin d'accéder à son substrat, DNMT1 pivote la cytosine à 180° hors de la double hélice d'ADN et procède alors au transfert de méthyle (Klimasauskas et al. 1994 ; figure 8). Récemment, la structure du complexe ADN-DNMT1 a été obtenue et révèle un mécanisme d'auto-inhibition dans le processus de maintenance de la méthylation de l'ADN, assurant uniquement la méthylation des dinucléotides hémiméthylés (Song et al. 2011).

Outre la région C-terminale, conférant à DNMT1 son activité catalytique, le domaine N-terminal est crucial et joue le rôle de régulateur.





**Figure 8. Structure 3D de mDNMT1.**

**A :** Structure 3D de mDNMT1 en complexe avec l'ADN.

**B :** Modèle du mécanisme d'auto-inhibition. Le domaine CXXC en liant spécifiquement les CpGs non méthylés, permet au « linker » CXXC-BAH1 d'empêcher l'interaction entre l'ADN et le site actif de DNMT1 et donc d'inhiber la réaction de méthylation. De plus, une interaction entre BAH2 et le domaine de reconnaissance TRD empêche ce dernier de s'insérer dans le grand sillon de l'ADN, déstabilisant ainsi le complexe.

(Source : Song et al. 2011).

**Table 2. Phénotype des souris KO.**

Gène	Phénotype des souris KO	Référence
<i>Dnmt1</i>	Létale embryonnaire	Li et al. 1992
<i>Dnmt2</i>	Viable et fertile Profil de méthylation de l'ADN normal	Okano et al. 1998
<i>Dnmt3a</i>	Létale 4 semaines après la naissance Défauts d'empreinte	Okano et al. 1999
<i>Dnmt3b</i>	Létale embryonnaire Hypométhylation des satellites mineurs	Okano et al. 1999
<i>Dnmt3l</i>	Viable mais stérile Défauts d'empreinte	Bourc'his et al. 2001
<i>Mecp2</i>	Létale 10 semaines après la naissance Problèmes neurologiques, similaire au syndrome de Rett	Chen et al. 2001
<i>Mbd1</i>	Viable et fertile Déficits de neurogenèse, comportement autistique	Zhao et al. 2003 Allan et al. 2008
<i>Mbd2</i>	Viable et fertile Comportement maternel altéré	Hendrich et al. 2001
<i>Mbd3</i>	Létale embryonnaire	Hendrich et al. 2001
<i>Mbd4</i>	Viable et fertile	Sansom et al. 2003
<i>Uhrf1</i>	Létale embryonnaire	Muto et al. 2002
<i>Uhrf2</i>	Indéterminé	
<i>Kaiso</i>	Viable et fertile	Prokhortchouk et al. 2006
<i>Zbtb4</i>	Viable et fertile	Non publié
<i>Zbtb38</i>	Indéterminé	

Le domaine N-terminal de DNMT1 a deux fonctions principales : la régulation allostérique de l'activité catalytique et la régulation de la stabilité de DNMT1 par des modifications post-traductionnelles comme la méthylation et la phosphorylation.

Il a été prédit que 3 à 20% de la méthylation de l'ADN serait perdue si les cellules ne dépendaient que de l'activité de DNMT1 (Ooi et Bestor 2008), ce qui suggère que d'autres facteurs sont nécessaires au recrutement et/ou à la régulation de DNMT1. Parmi les partenaires de DNMT1, PCNA et UHRF1 semblent les meilleurs candidats pour la régulation de l'activité catalytique. Cependant, la délétion du domaine d'interaction avec PCNA ne semble pas modifier l'activité de DNMT1 (Spada et al. 2007). Par ailleurs, PCNA ne possède pas d'affinité pour l'ADN hémiméthylé. Ensemble, ces résultats indiquent que PCNA n'est pas un recruteur de DNMT1. En revanche, UHRF1 se révèle avoir de l'affinité pour l'ADN hémiméthylé et sa délétion chez la souris présente un phénotype similaire à celui de DNMT1 (Muto et al. 2007). UHRF1 apparaît donc comme un facteur essentiel au mécanisme de maintenance de la méthylation de l'ADN (voir Partie I, chapitre 3 MBPs).

De récentes études ont montré que DNMT1 peut subir des modifications post-traductionnelles altérant son activité et sa stabilité. En fin de phase S, DNMT1 peut être méthylée par SET7, une lysine méthyltransférase identifiée originellement pour méthyler les histones. DNMT1 méthylée est alors moins stable et est dégradée par le protéasome (Estève et al. 2009). De plus, l'absence d'une déméthylase spécifique des lysines (LSD1) montre une augmentation de la dégradation de DNMT1 associée à une hypométhylation de l'ADN de H19 et des éléments IAP (Wang et al. 2009). DNMT1 peut également être phosphorylée par la kinase AKT1 ce qui permet d'accroître la demi-vie de DNMT1 comparé à la forme méthylée (Estève et al. 2011 ; Hervouet et al. 2010). En plus d'augmenter la stabilité de DNMT1, la phosphorylation induit une augmentation de son activité catalytique *in vitro* (Goyal et al. 2007). Il est aussi intéressant de noter que la méthylation et la phosphorylation sont mutuellement exclusives indiquant un mécanisme de régulation du niveau protéique de

DNMT1 au cours du cycle cellulaire (Estève et al. 2011). Enfin, une dernière modification a été dernièrement mise en évidence : la sumoylation. DNMT1 lie SUMO-1 et Ubc9 et devient sumoylée augmentant ainsi son interaction avec l'ADN et probablement son efficacité catalytique (Lee et al. 2009).

Outre sa fonction dans le maintien de la méthylation de l'ADN, DNMT1 possède une activité *de novo*. DNMT1 semble même avoir une activité plus importante que les enzymes *de novo* DNMT3a et DNMT3b *in vitro* (Okano et al. 1998). Plus récemment, il a été montré que DNMT1, en coopération avec DNMT3a, est capable de méthyler un ADN non modifié (Fatemi et al. 2002). Des données montrent que DNMT1 est capable d'initier la méthylation des CGIs suggérant ainsi son rôle dans le processus d'hyperméthylation des cellules cancéreuses (Jair et al. 2006).

### 2.1.3 DNMT2

Suite à la découverte de DNMT1 et de son activité de maintenance, de nombreux efforts ont été effectués afin d'identifier la DNMT *de novo* (figure 6). Des recherches d'analyse d'homologie de séquence ont alors été entreprises et ont mené à la découverte de DNMT2 (Okano et al. 1998 ; Yoder et Bestor 1998). Bien que DNMT2 soit la DNMT la plus conservée au cours de l'évolution, les analyses biochimiques de l'époque ne détectent aucune activité ADN méthyltransférase. De plus, aucun changement de méthylation de l'ADN n'est observé dans les cellules ES de souris (Okano et al. 1998).

Cependant, le séquençage au bisulfite du génome de *Drosophila* suggère que DNMT2 (unique DNMT présente chez la *Drosophila*) méthyle des cytosines sans spécificité de séquence (Lyko et al. 2000). Il a ensuite été montré que la protéine purifiée est capable de méthyler l'ADN *in vitro* (Hermann et al. 2003). Le rôle de DNMT2 dans l'inhibition de rétrotransposons a récemment été l'objet d'une étude chez la *Drosophila*. Les auteurs montrent que DNMT2 est requis pour l'inhibition de rétrotransposons ainsi que des répétitions sub-télomériques (Phalke et al. 2009). Cependant, ces travaux font l'objet de nombreuses controverses (Schaefer et al. 2010a).

Si DNMT2 ne semble pas avoir un rôle primordial dans la méthylation de l'ADN, il s'avère, en revanche, que DNMT2 possède une activité ARNt méthyltransférase (Goll et al. 2006). Si la fonction de la méthylation des ARNt n'est pas encore résolue, un début de réponse fut apporté par la démonstration que la méthylation par DNMT2 stabilise les ARNt chez la *Drosophile* (Schaefer et al. 2010b).

#### 2.1.4 DNMT3A et DNMT3B

DNMT3A et DNMT3B (figure 6) ont été identifiées en même temps que DNMT2 par homologie de séquence avec DNMT1 (Okano et al. 1998). A l'inverse de DNMT1, ces deux enzymes ne montrent pas de préférence pour l'ADN hémiméthylé. Il apparaît donc que ces enzymes jouent un rôle dans la mise en place des profils de méthylation de l'ADN au cours du développement plutôt que dans le maintien de ceux-ci. Les études d'expression montrent d'ailleurs que DNMT3A et DNMT3B sont fortement exprimées au stade blastocyste et également au cours de la différenciation des cellules germinales, deux étapes développementales au cours desquelles la méthylation de l'ADN *de novo* s'effectue. DNMT3A semble jouer un rôle particulièrement important dans les cellules germinales tandis que DNMT3B apparaît être nécessaire au cours du développement précoce.

Les expériences *in vitro* démontrent que DNMT3A et 3B ciblent préférentiellement les sites CpGs mais se lient également aux sites non CpG (Aoki et al. 2001). Il est intéressant de noter que la méthylation non CpGs est retrouvée dans les mêmes lignées exprimant fortement DNMT3A et 3B. *In vivo*, les cellules ES murines délétées pour *Dnmt3a* et *Dnmt3b* sont incapables de méthyliser un ADN proviral (Okano et al. 1999). De plus, la réintroduction de DNMT3A et 3B dans des cellules ES doubles mutantes, dépourvues de méthylation, permet de restaurer le profil de méthylation alors que DNMT1 en est incapable (Chen et al. 2003).

Par ailleurs, les cellules ES doubles mutantes présentent un faible taux de déméthylation à court terme alors qu'à plus long terme, le niveau de déméthylation se rapproche de celui des cellules mutantes pour *Dnmt1*. Ces

résultats indiquent que DNMT3A et 3B participent également au maintien de la méthylation (Chen et al. 2003). Sachant que l'activité *de novo* des DNMTs semble conduire à une hémiméthylation des CpGs (Dodge et al. 2002), il est nécessaire de méthyler symétriquement l'ADN avant le passage de la fourche de duplication. Dans le cas contraire, les cellules filles hériteraient d'un profil de méthylation différent. Cela nécessite donc un mécanisme de maintien indépendant de la machinerie de duplication. DNMT1, en reconnaissant l'ADN hémiméthylé, pourrait tenir ce rôle. Cependant, UHRF1, connue pour interagir directement avec DNMT3A et 3B (Jeong et al. 2009), pourrait également recruter les DNMTs *de novo* aux sites hémiméthylés expliquant ainsi la capacité de maintenance de DNMT3A et 3B (Meiling et al. 2009).

La structure de DNMT3A lié à son cofacteur DNMT3L suggère un nouveau modèle de méthylation *de novo* ayant lieu toutes les 8-10 pb. C'est en effet le cas pour les gènes maternels soumis à l'empreinte (Jia et al. 2007). En revanche, le modèle n'explique pas les profils de méthylation dans les régions méthylées riches en CpGs (Glass et al. 2009). Étant donné la périodicité a aussi été retrouvée dans le cas de la méthylation des CpGs associés aux nucléosomes. La signification biologique probable serait d'induire l'enroulement de l'ADN autour des nucléosomes (Chodavarapu et al. 2010). Compte tenu de la périodicité des dinucléotides CG autour des nucléosomes, il est plus probable de penser que cela permet l'accès des DNMTs aux sites à méthyler. La périodicité de méthylation observée serait donc indépendante de la structure du complexe DNMT3.

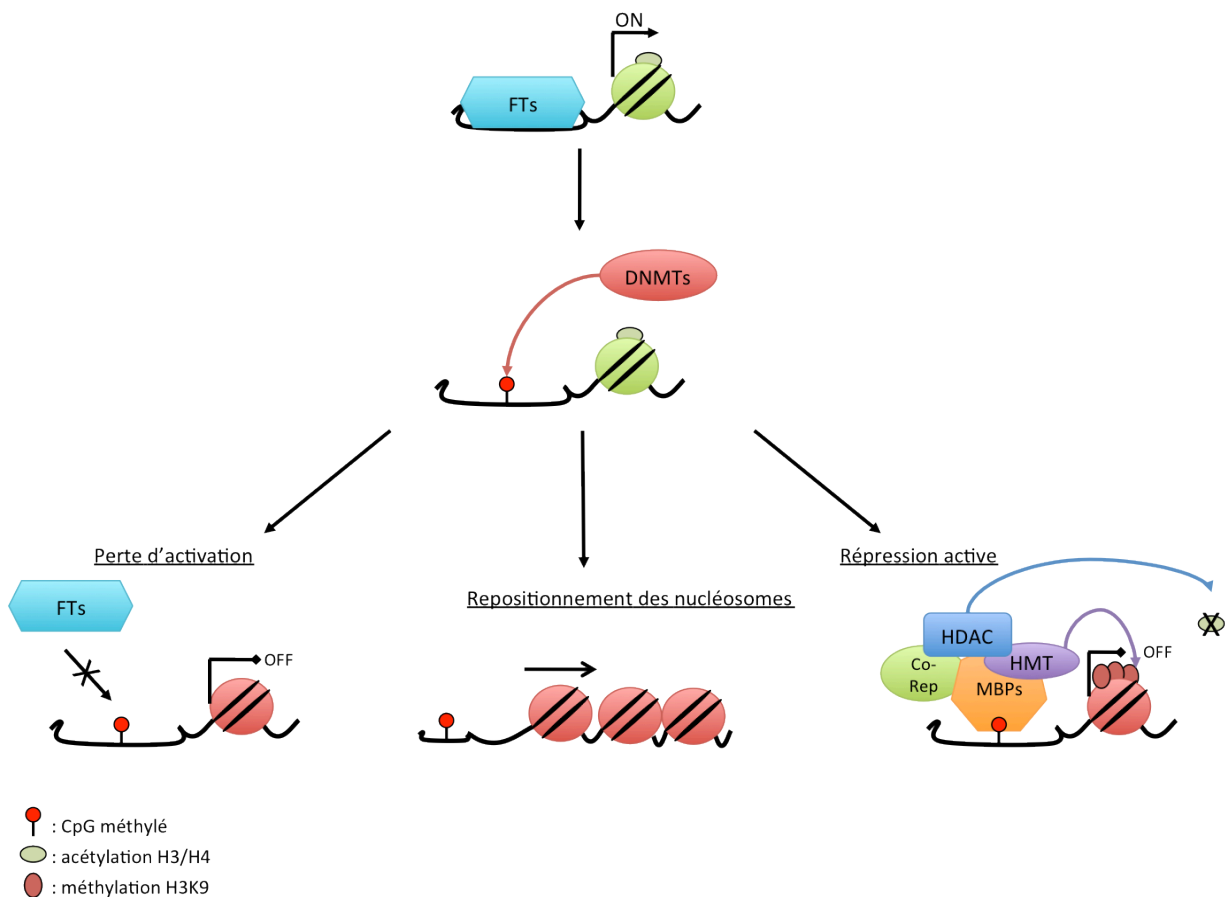
#### 2.1.5 DNMT3L

À partir de l'homologie du domaine N-terminal des DNMT3, une nouvelle ADN méthyltransférase a été identifiée : DNMT3-like (DNMT3L ; figure 6). Tout comme DNMT3A et 3B, DNMT3L possède un domaine C-terminal présentant des similarités avec les domaines catalytiques des DNMT3. Cependant, le domaine est tronqué et l'absence de résidus clés rend DNMT3L inactif en tant qu'ADN méthyltransférase. Les études biochimiques prouvent que DNMT3L lie faiblement l'ADN et qu'elle n'est pas capable de lier le SAM (Ooi et al. 2007).

DNMT3L semble avoir coévolué avec l’empreinte parentale suggérant un rôle de cette enzyme dans ce processus. L’invalidation du gène chez la souris tend à démontrer cette hypothèse. Les mâles sont stériles et les femelles présentent un défaut de méthylation des gènes maternels soumis à l’empreinte (Bourc’his et al. 2001).

L’interaction entre DNMT3L et DNMT3A et 3B permet la mise en place de la méthylation de l’ADN (Hata et al. 2002). DNMT3L se pose donc en tant que co-régulateur de DNMT3A et 3B. Il est intéressant de voir que DNMT3L permet de stimuler l’activité de DNMT3A et 3B (Chédin et al. 2002). La capacité des DNMT3 à lier l’ADN n’est pas modifiée mais l’affinité pour le SAM est augmentée rendant le complexe DNMT3-DNMT3L plus efficace (Suetake et al. 2004 ; Gowher et al. 2005 ; Kareta et al. 2006).

La capacité de DNMT3L à lier H3K4 non modifiée donne un début de réponse au ciblage de la méthylation *de novo*. Le domaine ADD de DNMT3L est incapable de lier H3K4me, marque active pour la transcription. DNMT3L ciblerait alors les histones non modifiées pour la mise en place de la méthylation de l’ADN. Cette anti-corrélation est d’ailleurs en adéquation avec la transcription des HCPs, sensiblement non méthylés et présentant un enrichissement pour H3K4me3.



**Figure 9. La méthylation des promoteurs peut inhiber la transcription par différents mécanismes.**

La méthylation peut empêcher la liaison de facteurs de transcription (FTs) et de co-facteurs nécessaires à l'activation des gènes. La méthylation de l'ADN peut également altérer la position des nucléosomes et donc influencer la transcription. Enfin, l'ADN méthylé peut recruter des complexes macromoléculaires spécifiques contenant des HDACs, des HMTs et autres co-répresseurs permettant la formation d'hétérochromatine.

## 2.2 Interprétation de la méthylation de l'ADN

Comme nous avons pu le voir au cours du chapitre 1, la méthylation de l'ADN est associée à la répression transcriptionnelle. Plusieurs mécanismes entrent en jeu dans cette inhibition (figure 9). Premièrement, la méthylation de la cytosine peut directement inhiber les interactions ADN-protéines et donc empêcher le recrutement des facteurs de transcription. C'est le cas, par exemple pour CREB (Iguchi-Ariga et al. 1989), c-Myc (Prendergast et al. 1991), E2F1 (Campanero et al. 2000) et UBF1 (Santoro et al. 2001). Deuxièmement, la méthylation des CpGs altère la conformation de la double hélice modifiant la position des nucléosomes. Ces changements régulent alors la liaison des facteurs de transcription et l'accès du complexe ARN polymérase (Segal et al. 2009). Finalement, certaines protéines reconnaissent directement la méthylation de l'ADN et induisent la formation d'hétérochromatine via des cofacteurs.

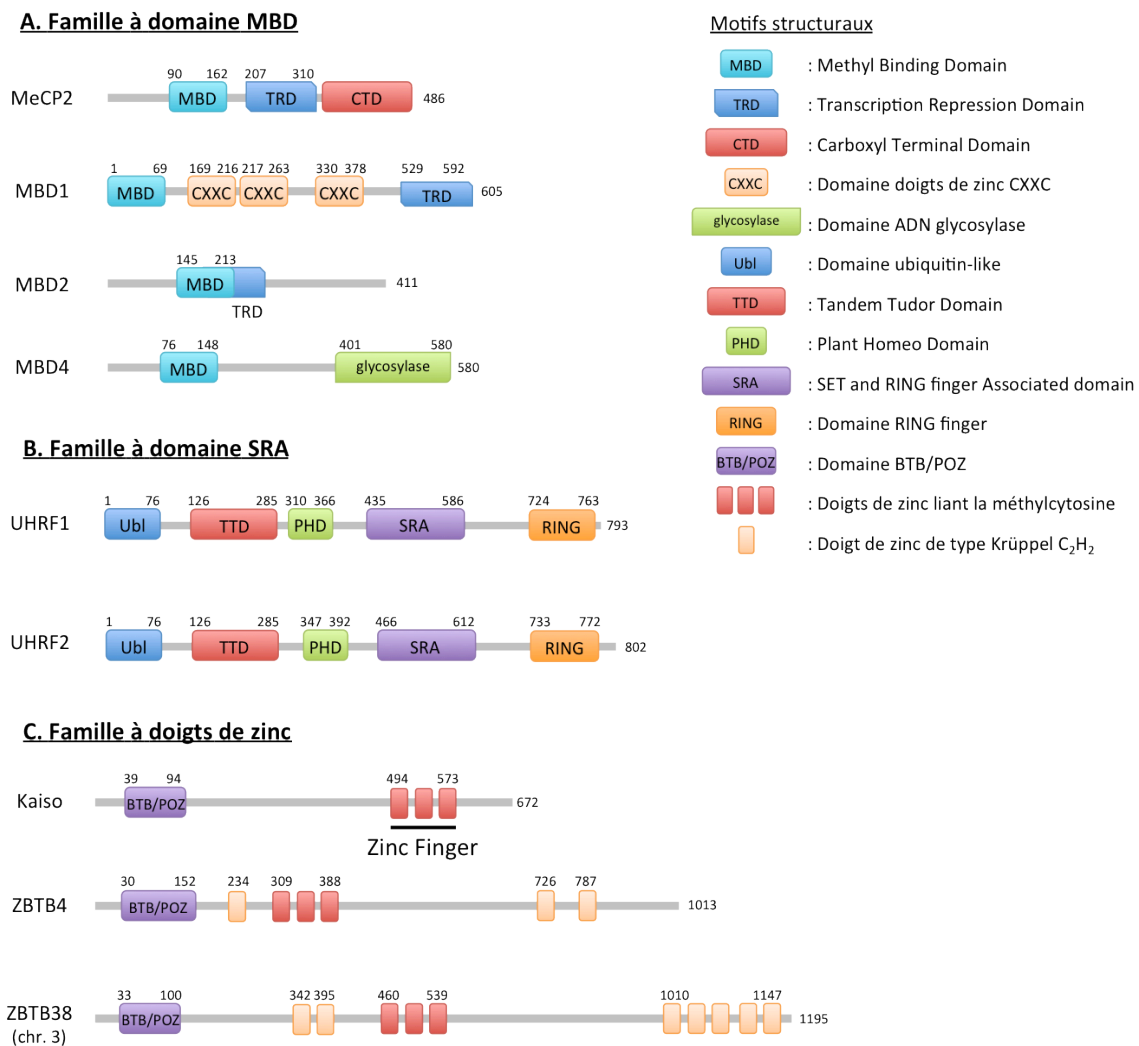
Trois familles de protéines sont connues pour lier l'ADN méthylé grâce à différents domaines structuraux. Les protéines de la famille MBD (Methyl Binding Domain) furent les premières identifiées, par le laboratoire d'Adrian Bird. Par la suite, ce même laboratoire montra qu'une protéine à doigt de zinc, Kaiso, lie également l'ADN méthylé. Notre équipe identifia alors deux protéines apparentées à Kaiso et ayant des propriétés voisines : ZBTB4 et ZBTB38. Enfin, tout récemment, deux protéines à domaine SRA (Set and Ring finger-Associated), UHRF1 et UHRF2, ont également été caractérisées pour se lier à l'ADN méthylé

### 2.2.1 Famille à domaine MBD

#### MeCP2

Les premières protéines liant l'ADN méthylé sont observées en 1989 par des expériences de retard sur gel à partir d'extraits nucléaires de foie de souris, incubés avec une sonde méthylée marquée radioactivement. Deux protéines sont ainsi identifiées et nommées MeCP1 et MeCP2 pour Methylated CpG-binding protein (Meehan et al. 1989). La purification de MeCP2 a été réalisée à partir d'extraits de cerveau de rat et a alors permis alors le clonage du gène *MeCP2* (Lewis et al. 1992 ; figure 10a).





**Figure 10. Les différentes familles de protéines se liant à l'ADN méthylé.**

**A :** Organisation des protéines MeCP2, MBD1, MBD2 et MBD4. Ces protéines partagent le même domaine methyl-CpG binding domain (MBD). Elles sont également caractérisées par des propriétés structurales spécifiques dont un domaine répresseur de la transcription (TRD), un domaine de compaction de la chromatine (CTD), un domaine de liaison non spécifique à l'ADN (CXXC), ou un domaine glycosylase. **B :** Domaines structuraux de UHRF1 et UHRF2. Les protéines sont très conservées et possèdent toutes deux un domaine ubiquitine-like (Ubl), un domaine Tudor (TTD), un domaine PHD, un domaine RING et un domaine SRA nécessaire à la liaison à l'ADN méthylé et hémiméthylé. **C :** Domaines structuraux de ZBTB4, kaiso et ZBTB38. Les protéines possèdent un domaine N-terminal BTB/POZ, de multiples doigts de zinc dont trois liant la méthylcytosine.

La caractérisation biochimique de MeCP2 montre qu'il existe un domaine de 90 acides aminés nécessaire et suffisant à la reconnaissance de l'ADN méthylé *in vitro*, le MBD (Nan et al. 1993), et un domaine de répression, le TRD (Transcriptional Repression Domain) (Nan et al. 1998). La résolution des structures cristallographique du domaine MBD montre que MeCP2 vient s'ancrer à l'ADN méthylé en formant une boucle en contact avec le grand sillon (Ho et al. 2008). Par ailleurs, MeCP2 présente une spécificité de liaison *in vitro*. En effet, cette protéine présente une plus grande affinité pour les CpGs méthylés suivis des bases A/T (Klose et al. 2005).

Dans les cellules de souris, MeCP2 est enrichie au niveau de structures hautement méthylées appelées chromocentres. Cette localisation est perdue dans les cellules ES *Dnmt1*<sup>-/-</sup> dépourvues de méthylation, confirmant que MeCP2 se lie aux CpGs méthylés *in vivo* (Nan et al. 1996).

Une caractéristique importante des MBPs est leur capacité à interagir avec des complexes modifiant la structure de la chromatine. Il a été décrit que le complexe inhibiteur SIN3A/HDAC1 ou HDAC2 interagit avec MeCP2 pour induire la répression (Nan et al. 1998). Par ailleurs, cette inhibition est réversible par traitement à la Trichostatine A (TSA), un inhibiteur de la déacétylation des histones, indiquant un lien étroit entre méthylation de l'ADN et modification des histones dans la mise en place des structures hétérochromatiniennes.

MeCP2 fait l'objet de nombreuses études car des mutations de ce gène sont responsables du syndrome de Rett caractérisé par un large panel de désordres neuropsychiatriques dont l'autisme, des troubles bipolaires, des déficits cognitifs, la schizophrénie... Les personnes atteintes de ce syndrome présentent généralement des troubles du langage, des retards psychomoteurs et des mouvements stéréotypés. Ce syndrome touche essentiellement les femmes, de par la localisation de MeCP2 sur le chromosome X, et il est létal chez les hommes (Amir et al. 1999). Les souris invalidées pour le gène *MeCP2* reproduisent de nombreux phénotypes du syndrome de Rett (table 2). Cependant, les mécanismes moléculaires précis

causant ces caractéristiques sont encore très peu compris. C'est pour cela que différentes délétions conditionnelles pour MeCP2 ont été réalisées dans des sous-types (sous-catégories) neuronaux (neurones dopaminergiques, neurones sérotoninergiques) mais les phénotypes décrits ne représentent pas l'ensemble des caractéristiques du syndrome de Rett. En revanche, la délétion de *MeCP2* dans les neurones GABAergiques récapitule pratiquement tous les troubles de ce syndrome de Rett démontrant ainsi l'importance de ces neurones dans cette pathologie (Chao et al. 2010).

Excepté le rôle dans le syndrome de Rett, MeCP2 est l'une des rares protéines dont nous connaissons des cibles. Elle a notamment été caractérisée en tant que répresseur transcriptionnel du gène *Bdnf* (Chen et al. 2003) et peut également réguler le gène *BRCA1* par la reconnaissance de la méthylation de l'ADN dans les tissus somatiques (Magdinier et al. 2000).

Cependant, une étude récente a démontré que MeCP2 était capable de réguler des milliers de gènes dont une majorité de façon active (voir Partie I, chapitre 4) (Chahrour et al. 2008). Cela suggère donc que MeCP2 a un rôle, non pas de régulateur spécifique d'un gène, mais plutôt comme régulateur global de la transcription. Afin de différencier ces deux fonctions, l'abondance et la distribution de MeCP2 ont été mesurées dans les neurones. Une molécule de MeCP2 est présente tous les deux nucléosomes confirmant le rôle de régulateur global de la transcription (Skene et al. 2010). Par ailleurs, il a été démontré que MeCP2 lie avec une plus forte affinité les nucléosomes méthylés plutôt que les nucléosomes non méthylés (Ghosh et al. 2010) et que son interaction avec l'ADN linker induit la compaction de l'ADN (Nikitina et al. 2007). La structure du complexe MeCP2-nucléosome récemment obtenue confirme ces résultats et montre que MeCP2 a potentiellement une fonction similaire à l'histone de « liaison » H1 (Yang et al. 2011).

A partir de la séquence du domaine MBD de MeCP2, quatre autres protéines possédant un domaine MBD ont été identifiées chez les mammifères : MBD1, MBD2, MBD3 et MBD4.

### MBD1

La structure de MBD1 en complexe avec l'ADN indique que le mécanisme de reconnaissance des CpGs méthylés est identique à celui de MeCP2 (Ohki et al. 2001) indiquant que les différents membres de la famille MBD, qui présentent une forte conservation, lient l'ADN méthylé de la même façon (figure 10a). Les études biochimiques montrent que, outre sa capacité à lier l'ADN méthylé double brin, MBD1 peut reconnaître un motif CGCG non méthylé par son domaine CXXC (Jorgensen et al. 2004). Récemment, des résultats *in vitro* ont permis de mettre en évidence que MBD1 présente une spécificité de séquence, dans la reconnaissance de la méthylation de l'ADN, en liant préférentiellement les motifs TmeCGCA ou TGmeCGCA (Clouaire et al. 2010).

Les souris mutées pour MBD1 présentent un phénotype autistique modéré et des défauts d'apprentissage (Zhao et al. 2003 ; Allan et al. 2008 ; table 2). L'absence de MBD1 empêche la différenciation et la neurogenèse des cellules neuronales progénitrices chez l'adulte, via la dérégulation d'une cible directe, *Fgf2* (Li et al. 2008). De plus, MBD1 régule directement l'expression du microARN miR-184 qui inhibe la différenciation des cellules progénitrices (Liu et al. 2010).

### MBD2

MBD2 est capable de lier l'ADN méthylé *in vitro* (Hendrich et al. 1998) et *in vivo* (Ng et al. 1999) et est associée à une activité répressive de la transcription, grâce à ses interactions avec les HDACs. A posteriori, MeCP1, identifiée en 1989, se révèle être le complexe composé de MBD2 et HDAC1. De la même manière que MeCP2, la répression par MBD2 est réversible par un traitement à la TSA (Ng et al. 1999).

D'un point de vue fonctionnel, la délétion de MBD2 chez la souris provoque une expression anormale des gènes pancréatiques dans le colon et indique que MBD2 intervient dans la régulation de gènes tissu-spécifiques (Berger et al. 2007 ; table 2). MBD2 régule également l'expression de l'interleukine 4 (IL4) dans les lymphocytes T (Hutchins et al. 2002) et réprime le promoteur méthylé de  $\rho$ -globine chez le poulet (Kransdorf et al.

2006) ainsi que le locus *BRCA1-NBR2* (Auriol et al. 2005), le gène *pS2/TFF1* (Chatagnon et al. 2010) et le gène *hTERT* (Chatagnon et al. 2009).

### MBD3

MBD3 partage 80% de similarité avec MBD2 mais l'absence de résidus clés dans son domaine MBD empêche toute liaison à l'ADN méthylé (figure 10a). Ironiquement, c'est la seule protéine de la famille MBD connue à ce jour pour être essentielle au cours du développement (Hendrich et al. 2001 ; table 3). Son rôle dans le complexe de remodelage de la chromatine Mi2-NuRD explique le phénotype léthal embryonnaire chez les souris *Mbd3<sup>-/-</sup>* (Zhang et al. 1999). Cependant, le rôle de MBD3 dans le complexe NuRD reste énigmatique. MBD3 interagit également avec MBD2 (Hendrich et al. 2001). Bien que le rôle de cette interaction ne soit pas entièrement compris, il semblerait que ce complexe puisse être recruté sur le promoteur méthylé de *Oct4* dans les cellules ES afin de contrôler la différenciation (Gu et al. 2011).

### MBD4

MBD4 est un membre à part de la famille des MBDs (figure 10a) car bien qu'elle soit capable de lier l'ADN méthylé et d'interagir avec les HDAC (Kondo et al. 2005), sa fonction principale semble être un senseur aux dommages à l'ADN. MBD4 possède un domaine glycosylase ainsi qu'une forte affinité pour les mésappariements mCpG•TpG issus de la déamination de la méthylcytosine (Hendrich et al. 1999). MBD4 contribuerait alors à maintenir l'intégrité du génome. De par son activité glycosylase, MBD4 excise la thymine et mobilise la machinerie de réparation afin de contrecarrer la mutagenèse spontanée de la méthylcytosine. Plus précisément, elle intervient dans le système de réparation des mésappariements (Mismatch Repair, MMR) par son interaction avec MutL Homolog 1 (MLH1) (Bellacosa et al. 1999). Des dommages à l'ADN induits par des UV vont entraîner le recrutement de MBD4, MLH1 et également DNMT1 afin d'éliminer la thymine et de restaurer la méthylcytosine (Mortusewicz et al. 2005 ; Ruzov et al. 2009). Ces observations sont corrélées au phénotype hypermutateur des souris *Mbd4<sup>-/-</sup>* qui présentent une augmentation de la transition C→T (Millar et al. 2002 ; Wong et al. 2002).

Par ailleurs, son interaction avec FADD (Fas-associated death domain protein) (Screaton et al. 2003) suggère un rôle de MBD4 dans l'apoptose. En effet, les souris *Mbd4*<sup>-/-</sup> présentent une réponse à l'apoptose diminuée lorsqu'elles sont soumises à un stress génotoxique (Sansom et al. 2003).

#### Les autres protéines à domaine MBD

Deux autres protéines possédant un domaine MBD ont été identifiées par similarité de séquence : MBD5 et MBD6 (Hendrich et 2003 ; Roloff et al. 2003). En dépit de leur association avec l'hétérochromatine qui requiert leur domaine MBD, ces protéines ne lient pas l'ADN méthylé (Laget et al. 2010).

BAZ2A (aussi appelée TIP5) et BAZ2B présentent également un domaine MBD et localisent dans des régions hétérochromatiniennes. Cependant, leur localisation est dépendante du complexe de remodelage de la chromatine auquel elles appartiennent : BAZ2A est un membre du complexe NoRC (Strohner et al. 2001) tandis que BAZ2B fait partie du complexe SNF2h.

Enfin, SETDB1 et SETDB2, deux histones méthyltransférases, contiennent également un domaine MBD qui n'a pas été rapporté comme étant capable de lier l'ADN méthylé.

#### **2.2.2 Famille à domaine SRA**

##### UHRF1

UHRF1 fait l'objet d'un grand intérêt depuis plus de 10 ans et présente un grand nombre de fonctions telles que le contrôle de la prolifération et du cycle cellulaire, la régulation de l'héritabilité des marques épigénétiques ou encore le maintien de l'intégrité du génome.

Historiquement, UHRF1 est caractérisée chez la souris sous le nom de Np95 pour apparaître spécifiquement en phase S dans des thymocytes normaux tandis qu'il est constitutivement exprimé dans une lignée cellulaire de lymphome. Par la suite, le gène fut cloné à partir d'ADNc et la séquence protéique révéla des interactions potentielles avec des acteurs du cycle cellulaire tels que la cycline A/E et la protéine Rb retinoblastoma,

suggérant un rôle de Np95 dans la prolifération et le cycle cellulaire (Fujimori et al. 1998). Cette hypothèse a été confirmée par la colocalisation de Np95 avec PcnA au niveau de la fourche de duplication au cours de la phase S (Uemura et al. 2000).

En parallèle, UHRF1 fut identifiée chez l'Homme par un crible de simple hybride pour interagir et réguler positivement le promoteur de la *Topoisomérase IIalpha* (Hopfner et al. 2000 ; Bronner et al. 2002 ; Hopfner et al. 2002). UHRF1 promeut également l'entrée en phase en S en réprimant le gène *Rb* (Jeanblanc et al. 2005).

L'ensemble de ces résultats souligne le rôle de UHRF1 dans la régulation du cycle cellulaire et plus précisément dans la transition G1/S par la répression de *p53* et *p21* (Mousli et al. 2003 ; Arima et al. 2004). Pour cela, UHRF1 est régulée de façon post-traductionnelle par phosphorylation par les kinases PKA et CK2, requises pour la prolifération cellulaire (Trotzler et al. 2004 ; Bronner et al. 2004).

L'invalidation du gène *mUhrf1* chez la souris est létale à l'état embryonnaire au stade 10,5 (dpc) démontrant le rôle essentiel de UHRF1 (Muto et al. 2002 ; table 2), tandis que les cellules ES *mUhrf1*<sup>-/-</sup> présentent un phénotype similaire à celui des mutants *Dnmt1*<sup>-/-</sup>, avec une hypométhylation globale du génome (Bostick et al. 2007 ; Sharif et al. 2007). Ces observations peuvent être expliquées par les partenaires de UHRF1, sa capacité de liaison à l'ADN et sa localisation au cours du cycle cellulaire. Premièrement, il a été montré que UHRF1 interagit directement avec DNMT1 (Bostick et al. 2007 ; Sharif et al. 2007 ; Achour et al. 2009). Deuxièmement, UHRF1 a été identifiée comme étant une protéine liant l'ADN méthylé *in vitro* grâce à son domaine SRA (Unoki et al. 2004 ; figure 10b) et présentant une plus grande affinité pour les sites hémiméthylés (Bostick et al. 2007). D'ailleurs, l'obtention de la structure a permis de mettre en évidence un mécanisme de reconnaissance semblable aux ADN méthyltransférases (Arita et al. 2008 ; Hashimoto et al. 2008 et Avvakumov et al. 2008). Le domaine SRA entre en contact à la fois avec le petit et le grand sillon de l'ADN et provoque le pivotement de la méthylcytosine hors de la double hélice permettant ainsi

l'accès à DNMT1 à la cytosine présente sur l'autre brin. Finalement, UHRF1 colocalise avec DNMT1 et PCNA en fin de phase S, au niveau des fourches de réplication tardives (Sharif et al. 2007 ; Bostick et al. 2007).

L'ensemble de ces résultats a permis d'éclaircir le mécanisme de transmission épigénétique de la méthylation de l'ADN au cours du cycle cellulaire en suggérant le rôle central de UHRF1 dans le recrutement de DNMT1 au niveau des sites hémiméthylés. De plus, l'interaction entre UHRF1-DNMT3A/3B (Meilinger et al. 2009) permet le maintien de la méthylation de l'ADN en dehors de la duplication (chapitre 2.1.4)

Par ailleurs, UHRF1 fait également le lien avec les modifications des histones. D'une part, UHRF1 participe au recrutement de modificateurs de la chromatine via ses interactions avec HDAC1 (Unoki et al. 2004), avec G9a, une histone méthyltransférase (Kim et al. 2009) et avec Tip60, une histone acétyltransférase (Achour et al. 2009). D'autre part, UHRF1 reconnaît les modifications des histones. Le domaine PHD associé au domaine SRA semble définir une liaison spécifique aux histones H3 méthylés sur leur lysine 9 (H3K9me) (Karagianni et al. 2008). Enfin, le domaine Tudor (TTD) de UHRF1 se lie aussi à H3K9me3 (Rottach et al. 2009) ainsi qu'à H3K4 non modifié (Nady et al. 2011). Ces données montrent le lien étroit entre méthylation de l'ADN et méthylation des histones. UHRF1 apparaît donc comme un facteur essentiel dans la communication entre les différentes marques épigénétiques.

En conséquence de son rôle dans le maintien de la méthylation de l'ADN, UHRF1 s'avère réguler les éléments répétés du génome ainsi que les gènes soumis à l'empreinte. En effet, la délétion de *mUHRF1* dans les cellules ES va induire la réexpression des rétrotransposons IAP, LINE-1 et SINE ainsi que des gènes « imprimés » *Igf2-H19*, *Kcnq1ot1*, et *Dlk1-Gtl2* provoqué par un défaut de méthylation de l'ADN (Bostick et al. 2007 ; Sharif et al. 2007).

UHRF1 semble jouer un rôle dans l'établissement correct des structures hétérochromatiennes par la reconnaissance des régions péricentromériques,



enrichies en H3K9me3. Le recrutement de HDAC1, DNMT1 et certainement d'autres enzymes modifiant la chromatine permettrait la réorganisation de ces structures (Papait et al. 2007 ; Papait et al. 2008 ; Karagianni et al. 2008).

UHRF1 intervient également dans de nombreux cancers où son expression est constitutive au cours du cycle. De plus, UHRF1 reconnaît les promoteurs méthylés des suppresseurs de tumeurs *p16<sup>INK4</sup>* et *p14<sup>ARF</sup>* et induit leur répression via le recrutement de HDAC1 dans les cellules cancéreuses (Unoki et al. 2004). UHRF1 régule également l'expression du pro-angiogène *VEGF* par interaction avec DNMT1 (Achour et al. 2008). UHRF1 inhibe également *p21* par son interaction avec l'histone méthyltransférase G9a (Kim et al. 2009). Ces résultats indiquent que UHRF1 se comporte comme un proto-oncogène.

## UHRF2

UHRF2 a été identifiée par un crible de double hybride pour interagir avec la protéine PCNP (PEST-containing nuclear protein) dont la fonction n'est pas clairement établie (Mori et al. 2002). UHRF2 est également capable de lier l'ADN méthylé (Unoki et al. 2004 ; figure 10b) et montre une préférence pour les sites hémiméthylés (Pichler et al. 2011). Cependant, la fonction de UHRF2 semble indépendante de la méthylation de l'ADN et ce d'après deux raisons majeures. Premièrement, UHRF2 est incapable de restaurer les profils de méthylation des cellules *mUhrf1<sup>-/-</sup>* (table2). Deuxièmement, la localisation de UHRF2 dans le noyau dépend de H3K9me3 et non de la méthylation de l'ADN.

UHRF2, comme UHRF1, a un rôle dans le contrôle du cycle cellulaire. Mais à la différence de UHRF1, la surexpression de UHRF2 induit l'arrêt des cellules en G1 (Li et al. 2004). Alors que UHRF1 est exprimée dans les cellules souches pluripotentes, UHRF2 est transcrite au cours de la différenciation et dans les tissus différenciés (Pichler et al. 2011). Ces observations indiquent un rôle antagoniste de UHRF2 et suggèrent une boucle de rétro-contrôle entre UHRF1 et UHRF2 permettant de réguler l'activité du cycle cellulaire et de la différenciation.

### 2.2.3 Famille de protéines à doigts de Zinc

#### Kaiso

Une protéine à doigts de zinc, nommée Kaiso, a été caractérisée pour lier l'ADN méthylé (figure 10c). Des expériences avec un plasmide rapporteur méthylé montrent que Kaiso est capable d'induire une répression transcriptionnelle (Prokhortchouk et al. 2001). Contrairement aux autres MBPs, Kaiso (ainsi que les autres protéines à doigts de zinc) ne lie pas seulement les CpGs méthylés mais également une séquence spécifique non méthylée appelée KBS (Kaiso binding site) (Daniel et al. 2002). Cette dernière contient un TpG dont la thymine présente une structure analogue à celle d'une méthylcytosine, expliquant probablement la capacité de Kaiso à lier ces deux types de séquences.

Kaiso fait partie de la famille des protéines ZBTB qui possèdent un domaine d'interaction protéique BTB/POZ (Perez-torrado et al. 2006) ainsi que des doigts de zinc de type Krüppel-like C2H2. Kaiso est associée à des protéines qui remodelent la chromatine et recrute le complexe répresseur NCoR via son domaine BTB/POZ (Yoon et al. 2003).

Kaiso a été étudiée principalement dans deux systèmes : chez les mammifères et chez le Xénope. La délétion de *xKaiso* chez le Xénope provoque une transcription précoce de nombreux gènes conduisant à l'apoptose et à un arrêt développemental (Ruzov et al. 2004). Ce phénotype est similaire à celui obtenu lorsque *xDnmt1* est invalidé chez le Xénope indiquant que la fonction de xKaiso dans le développement est dépendante de la méthylation. Par ailleurs, xKaiso s'avère essentielle pour la gastrulation au cours du développement en régulant la voie Wnt (Kim et al. 2004 ; Park et al. 2005). Cependant, la mutation chez la souris est viable et les individus sont fertiles et ont une durée de vie normale (Prokhortchouk et al. 2006 ; table 2).

Récemment, il a été montré que Kaiso localise au niveau des centrosomes et du fuseau de microtubules au cours de la mitose suggérant un lien avec le cycle cellulaire et la prolifération (Soubry et al. 2010).

**Table 3. Séquences cibles des MBPs.**

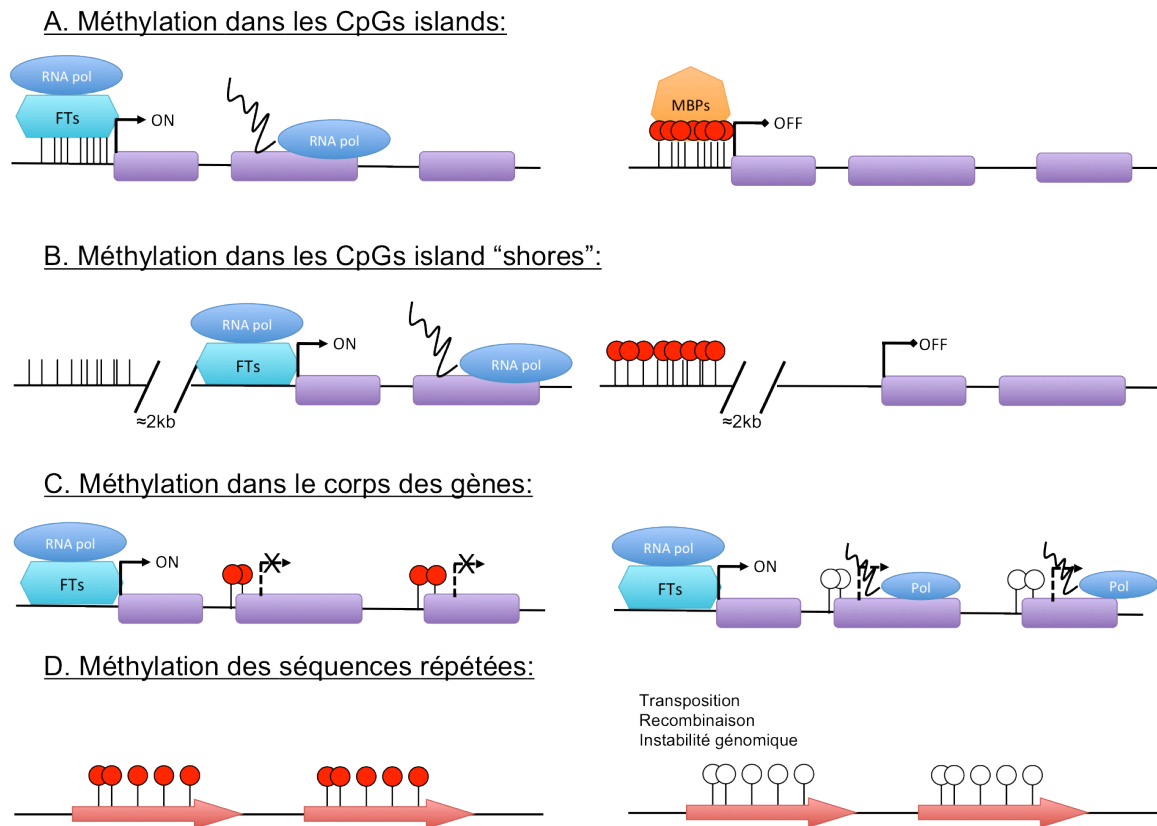
Protéine	Séquence cible	Référence
MeCP2	Motif A/T flanquant le méthyle CpG	Klose et al. 2005
MBD1	TMGCA et TGMGCA	Clouaire et al. 2010
MBD2	Pas de spécificité connue	Klose et al. 2005
MBD3	Ne lie pas la méthylation de l'ADN	Hendrich et Bird 1998
MBD4	mCpG•TpG	Hendrich et al. 1999
MBD5	Ne lie pas la méthylation de l'ADN	Laget et al. 2010
MBD6	Ne lie pas la méthylation de l'ADN	Laget et al. 2010
UHRF1	CpG hémiméthylé	Unoki et al. 2004
UHRF2	CpG hémiméthylé	Unoki et al. 2004
Kaiso	Deux méthyles CpGs consécutifs Séquence non méthylée CTGCNA	Prokhortchouk et al. 2001
ZBTB4	Séquence méthylée $^C/_A\text{MGC}^C/_T\text{AT}$ Séquence non méthylée $C^C/_T\text{GCCATC}$	Sasai et al. 2010
ZBTB38	Séquence méthylée $^C/_A\text{MGC}^C/_T\text{AT}$	Sasai et al. 2010

### ZBTB4 et ZBTB38

Notre laboratoire a identifié par homologie de séquence avec Kaiso deux autres protéines capables de lier l'ADN méthylé : ZBTB4 et ZBTB38 (Filion et al. 2006 ; figure 10c). Elles présentent les mêmes caractéristiques que Kaiso : liaison aux CpGs méthylés et à une séquence non méthylée. Toutefois, leur liaison à l'ADN méthylé a lieu au niveau d'une séquence méthylée spécifique suggérant alors que leur fonction n'est pas redondante avec les autres MBPs (Sasai et al. 2010).

ZBTB4 est dérégulée dans les cancers et semble jouer le rôle de suppresseur de tumeur. ZBTB4 inhibe la transcription de p21, favorisant ainsi l'apoptose en cas de dommages à l'ADN, tandis que son absence conduit à un arrêt du cycle cellulaire (Weber et al. 2008). De plus, la kinase HIPK2, qui contrôle la prolifération cellulaire, régule négativement ZBTB4 par phosphorylation en cas de dommages à l'ADN (Yamada et al. 2009). En revanche, les souris mutantes pour *Zbtb4* ont récemment été obtenues au sein du laboratoire et sont viables.

ZBTB38 induit la répression via le recrutement d'un complexe répresseur CtBP1/HDAC au niveau des péricentromères (Sasai et al. 2005). ZBTB38 joue également un rôle dans le contrôle du cycle cellulaire et de la réponse aux dommages à l'ADN via la régulation de MCM10 (non publié).



**Figure 11. Distribution de la méthylation des CpGs.**

**A :** Les îlots CpGs sont généralement non méthylés et permettent la transcription. Certains promoteurs de gènes spécifiques des tissus deviennent hyperméthylés au cours de la différenciation et sont alors transcriptionnellement inactifs. **B :** L'hyperméthylation des îlots CpGs situés en amont du promoteur est associée à la répression transcriptionnelle. La méthylation des îlots CpGs « shores » correspond aux régions différenciellement méthylées entre les tissus **C :** La méthylation dans le corps des gènes est associée à l'expression des gènes et permettrait d'empêcher l'initiation interne de la transcription. Un rôle dans le processus d'épissage a également été proposé. **D :** La méthylation des séquences répétées permet le maintien de l'intégrité du génome.

Source : Portela et Esteller, 2010.

## Chapitre 3 : rôle de la méthylation de l'ADN

### 3.1 Distribution de la méthylation de l'ADN

Afin de mieux comprendre les fonctions et la dynamique de méthylation de l'ADN, il est nécessaire d'avoir une vision globale à l'échelle du génome. Les récentes avancées technologiques nous permettent désormais d'obtenir des profils de méthylation de différents génomes, de différents tissus et de différentes maladies. Ces cartes pourront nous renseigner sur la mise en place de la méthylation au cours du développement et de la différenciation mais apporteront également des éclaircissements sur la progression des maladies.

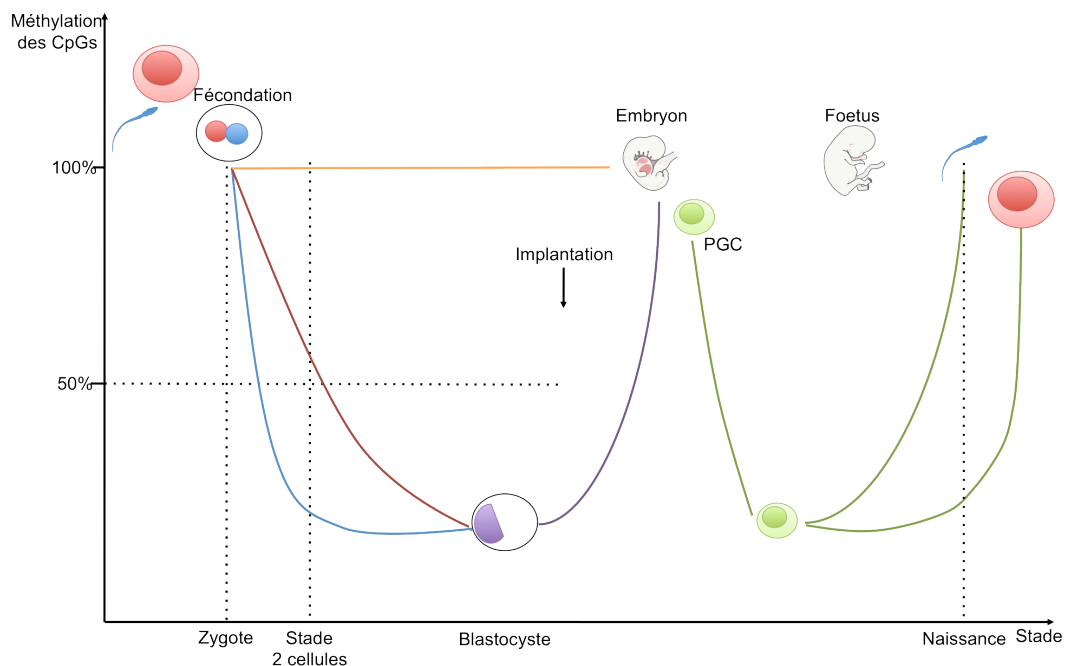
De manière globale, les ADN méthylomes obtenus montrent un important niveau de méthylation dans les cellules souches embryonnaires. Puis le niveau global tend à diminuer au cours de la différenciation (Laurent et al. 2010). Une caractéristique intéressante est le niveau de méthylation non-CpG retrouvé dans les cellules ES et notamment la méthylation CpA, suggérant un rôle potentiel de cette modification dans le maintien de la pluripotence (Lister et al. 2010).

Ensuite, la méthylation de l'ADN est plus importante dans le corps des gènes transcrits et diminue lorsque l'on s'éloigne du site de transcription (TSS) (Zhang et al. 2006 ; Ball et al. 2009). De façon plus précise, la méthylation est enrichie dans les exons plutôt que dans les introns avec un pic au niveau du 5' de l'exon et une réaugmentation dans le 3' de l'intron (Laurent et al. 2010 ; figure 11). Il est probable que la méthylation joue ici un rôle dans la régulation de l'épissage des ARNm.

Les profils réalisés dans différentes lignées cellulaires permettent une analyse comparative et démontrent l'existence de régions différentiellement méthylées (RDM) entre les tissus (Weber et al. 2007 ; Yagi et al. 2008). Ces RDMs sont retrouvées principalement au niveau des CGIs des promoteurs et des régions intragéniques. Comme discuté précédemment, la méthylation des CGIs intragéniques est anti-corrélée à H3K4me3, ce qui suggère un mécanisme possible de régulation des promoteurs alternatifs afin de

maintenir la spécificité des tissus (Maunakea et al. 2010). Les RDMs identifiées dans ces études correspondent à des gènes soumis à l’empreinte, des gènes de pluripotences ou nécessaire à la différenciation. Confirmant également des données connues depuis longtemps, les profils de méthylation montrent une hypométhylation globale du génome et une hyperméthylation des gènes suppresseurs de tumeurs dans les cancers. Mais de façon plus intéressante, les miARNs font également partis de RDMs et sont inhibés dans les cancers (Lujambio et al. 2007).

L’autre avantage de ces nouvelles approches globales est l’identification de méthylation allèle-spécifique (MAS). Il y a non seulement les gènes soumis à l’empreinte mais aussi environ 10% des gènes humains régulés de cette manière (Zhang et al. 2009).



**Figure 12. La méthylation de l’ADN au cours du développement.**

La ligne bleue correspond à la déméthylation active du génome paternel. La ligne rouge représente la déméthylation passive du génome maternel. Les gènes soumis à l’empreinte échappent à la déméthylation globale du génome (ligne orange). Au stade blastocyste, le génome est reprogrammé et les profils de méthylation sont mis en place (ligne violette). Une autre vague de reprogrammation a lieu dans les cellules germinales primordiales (PGC, ligne verte). Les gènes soumis à l’empreinte sont alors déméthylés puis reméthylés afin d’établir les profils de méthylation dépendant de l’origine parentale.

## 3.2 La méthylation de l'ADN au cours du développement

En dépit de son rôle relativement stable dans la répression à long terme, la méthylation de l'ADN apparaît dynamique et flexible au cours du développement (figure 12). Deux vagues distinctes de reprogrammation du génome sont ainsi effectuées et caractérisées par la déméthylation-reméthylation de l'ADN. La première a lieu dans les œufs fertilisés tandis que la seconde se produit dans les cellules germinales primordiales (PGC).

### 3.2.1 Reprogrammation de la méthylation de l'ADN des embryons pré-implantatoires

Chez les mammifères et plus précisément chez la souris, d'importants changements épigénétiques sont observés à partir de la fertilisation jusqu'à l'implantation de l'embryon. Durant cette période, le pronucléus paternel subit une déméthylation globale de son génome dans les 4 à 8 heures suivant la fertilisation, avant même que la première division cellulaire est eu lieu (Mayer et al. 2000 ; Oswald et al. 2000). La rapidité du processus de déméthylation suggère alors qu'il s'agit d'un mécanisme actif indépendant de la duplication. Par ailleurs, le traitement des zygotes avec un inhibiteur de la duplication tel que l'aphidicoline n'empêche pas la déméthylation du génome paternel (Mayer et al. 2000).

Bien que les études d'immunomarquages aient montré le caractère global de la déméthylation, les séquençages au bisulfite indiquent que certaines régions sont résistantes à la reprogrammation. C'est notamment le cas pour les gènes soumis à l'empreinte (Olek et al. 1997), les régions répétées centro- et péricentromériques (Rougier et al. 1998 ; Santos et al. 2002) et certains éléments transposables comme le rétrotransposon IAP (Intracisternal A-particle) (Lane et al. 2003). Mais récemment, il a été montré que certains gènes, habituellement exprimés dans la lignée germinale mâle, échappent également à cette déméthylation (Borgel et al. 2010).

A l'inverse, le génome maternel semble protégé de cette vague de déméthylation active par des mécanismes encore inconnus. Toutefois le génome maternel va perdre progressivement la méthylation de l'ADN par



un processus passif de dilution de la méthylation au cours des divisions cellulaires. En effet, l'isoforme de DNMT1 présent dans l'oocyte est séquestré au sein du cytoplasme jusqu'au stage 8 cellules (Carlson et al. 1992) empêchant le processus de maintenance de la méthylation de l'ADN.

Ensuite, au cours de la transition du blastocyste à l'épiblaste post-implantatoire, la méthylation est rétablie et cible notamment les gènes de la lignée germinale ainsi que les gènes tissus-spécifiques tels que les gènes hématopoïétiques (Borgel et al. 2010). De façon plus général, les promoteurs avec un taux faible (LCP, Low CpG Promoter) ou intermédiaire de CpGs (ICP, Intermediate CpG Promoter) sont hyperméthylés tandis que les promoteurs avec un taux élevé de CpGs (HCP, High CpG promoter) restent globalement hypométhylé (Weber et al. 2007 ; Meissner et al. 2008).

### **3.2.2 Reprogrammation de la méthylation de l'ADN des PGCs**

Les cellules postérieures de l'épiblaste se différencient en PGCs dans les embryons de 7,5 jours (E7,5) à partir de signaux provenant de l'ectoderme extraembryonnaire et de l'endoderme viscéral. Pendant la gastrulation, entre E8,5 et E11,5, les cellules migrent vers la crête génitale et subissent une nouvelle vague de déméthylation (Hajkova et al. 2002). Cette reprogrammation affecte l'ensemble du génome dont les gènes soumis à l'empreinte.

Le processus de déméthylation semble mettre en jeu des mécanismes actifs sachant que les PGCs vont se diviser à plusieurs reprises en présence de DNMT1. Le génome des futurs gamètes est alors reméthylé, notamment sur les gènes soumis à l'empreinte afin d'établir les profils de méthylation dépendant de l'origine parentale. Dans les cellules germinales mâles, la reméthylation des éléments répétés débute vers E15,5 se termine 3 jours après la naissance. Dans le cas des gamètes femelles, le processus est moins clair mais l'empreinte maternelle est rétablie après la naissance (Bourc'his et al. 2001).

### 3.3 Contrôle de l'expression de certains gènes tissu-spécifiques

Les tissus somatiques possèdent des profils d'expression uniques correspondant aux lignées cellulaires différenciées. Rapidement, il a été proposé que la méthylation de l'ADN soit le mécanisme responsable de la régulation de l'expression des gènes tissu-spécifiques. C'est dans cette optique qu'un grand nombre d'études a été réalisé afin de corrélérer l'expression des gènes avec les profils de méthylation de l'ADN. Parmi les nombreux exemples existants, on peut citer trois cas intéressants où la méthylation de l'ADN est étroitement liée à l'expression des gènes tissu-spécifiques. Dès 2002, l'équipe de Dommann montre que le gène *SERPINB5* codant pour la Maspin est exprimé dans les tissus épithéliaux de seins, de prostate et dans les kératinocytes de peau mais pas dans les fibroblastes, les lymphocytes, le cœur ou encore les reins (Futscher et al. 2002). Ces niveaux d'expression sont parfaitement corrélés à la méthylation des CGIs localisés dans les régions promotrices. Ces résultats ont été les premiers à supporter le rôle de la méthylation des cytosines dans l'établissement et le maintien de l'expression des gènes de façon spécifique au type cellulaire. La preuve du rôle de la méthylation de l'ADN dans la régulation des gènes tissu-spécifiques est apportée par l'étude d'expression des gènes de la lignée germinale. En effet, les gènes *Mvh*, *Daz1* et *Scp3* sont déméthylés dans les cellules germinales de souris et transcriptionnellement actifs. Au contraire, dans les cellules somatiques, ces gènes sont méthylés et inactifs. Mais surtout, les embryons de souris invalidés pour *Dnmt1* présentent une expression précoce des gènes de la lignée germinale suggérant le rôle crucial de la méthylation de l'ADN (Maatouk et al. 2006). Enfin, une analyse de la méthylation des CGIs dans le sang, le cerveau, le muscle et la rate chez l'Homme montre un enrichissement au niveau de gènes essentiels pour le développement tels que les membres de la famille *HOX* et *PAX* (Illingworth et al. 2008). Par ailleurs cette même étude démontre que la méthylation des CGIs est majoritairement retrouvée sur des séquences inter- et intragénique suggérant que les régions différentiellement méthylées (RDM) entre les tissus ne sont pas localisées sur les promoteurs. Les études à l'échelle du génome tendent à confirmer cette hypothèse et montrent que la majorité des RDMs se trouve à environ 2 kb des promoteurs, dans les CGIs « shores » (Irizarry et al. 2009 ; Doi et al. 2009).

### 3.4 Contrôle de l'expression des éléments répétés

Le génome humain est constitué pour moitié de séquences répétées (Mills et al. 2007). Il en existe deux types: les séquences répétées en tandem correspondant à l'ADN satellite principalement retrouvé au niveau des centromères et des télomères, et les séquences répétées dispersées correspondant aux éléments transposables répartis sur l'ensemble du génome. La plupart de ces séquences sont méthylées, et elles contiennent la majorité des cytosines méthylées du génome. En revanche, la méthylation de ces deux types de séquences a vraisemblablement des rôles différents.

#### 3.4.1 Maintien de l'intégrité du génome

Les éléments transposables sont répartis en deux catégories : les transposons et les rétrotransposons possédant une activité transcriptase reverse. La transposition de ces éléments représente une menace de l'intégrité des génomes de mammifères. Premièrement, leur activité est source de mutation par intégration dans les gènes. Deuxièmement, leur nombre et leur homologie peuvent permettre à des événements de recombinaison entre chromosomes non homologues d'avoir lieu.

La méthylation de ces éléments répétés indique un rôle de la méthylation de l'ADN dans la défense des génomes contre ces éléments. La méthylation des rétrotransposons inhiberait ainsi la transcription et donc la réplication de ces éléments. Tandis que la méthylation des transposons pourrait empêcher les événements de recombinaison d'avoir lieu. Cette hypothèse est appuyée par les observations effectuées dans les cellules ES mutées pour *Dnmt1* qui présentent une instabilité génomique (Chen et al. 1998). De plus, les éléments IAP sont déméthylés et réactivés dans les embryons *Dnmt1*<sup>-/-</sup> (Walsh et al. 1998). Par ailleurs, la méthylation des éléments transposables a lieu au cours de la gamétogenèse et nécessite DNMT3L. En effet, la délétion de *Dnmt3l* provoque une hypométhylation et une activité de transposition accrue (Bourc'his et Bestor 2004 ; Webster et al. 2005).

Bien que les mécanismes de méthylation des éléments transposables ne soient pas totalement résolus, il apparaît clairement que des petits ARNs de 25-30 nt appelés piARNs permettent l'établissement de la méthylation et le contrôle des séquences répétées (Aravin et al. 2007). L'analyse de la méthylation de l'ADN et de l'expression des piARNs dans des cellules germinales fœtales invalidées pour un gène de la famille *Argonaute* : *Piwi* montre une hypométhylation des éléments LINE-1 et IAP associée à une réduction de l'expression des piARNs (Kuramochi-Miyagawa et al. 2008). Ces résultats suggèrent que les piARNs sont requis dans la méthylation *de novo* des éléments transposables dans les cellules germinales (Aravin et al. 2008).

#### 3.4.2 Rôle dans l'établissement des structures hétérochromatiques

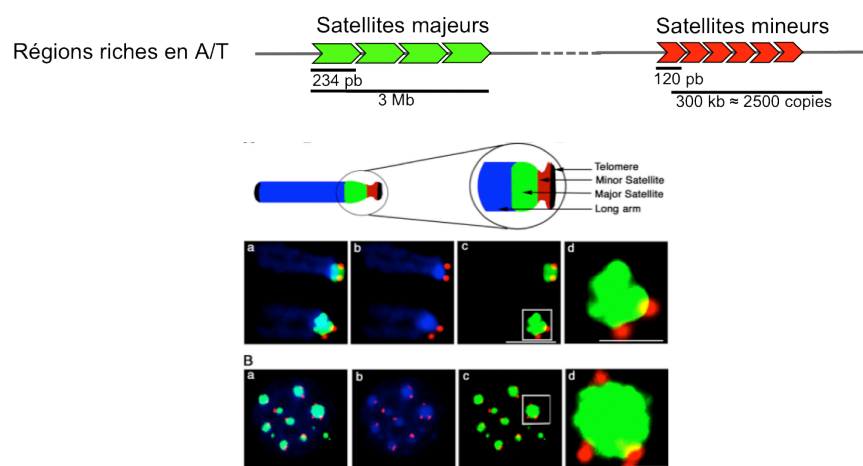
Outre les éléments transposables, le génome contient d'autres types de séquences répétées qui sont responsables de la mise en place des structures hétérochromatiniennes. C'est le cas, chez l'Homme et la souris, des régions péricentriques qui permettent de définir une architecture correcte des centromères nécessaire à la cohésion des chromatides sœurs et à la ségrégation des chromosomes au cours de la mitose et de la méiose (Probst et Almouzni 2011).

Chez la souris, les centromères et péricentromères représentent la majorité des domaines d'hétérochromatine du génome. Les centromères sont composés de nombreuses répétitions riches en A-T de 120 pb s'étalant sur près de 600 kb et appelées satellites mineurs (Choo et al. 1997), tandis que les péricentromères sont des séquences répétées adjacentes aux centromères appelées satellites majeurs. Ces répétitions sont également riches en A-T et le motif est long de 234 pb réparties sur plus de 2 Mb (figure 13 ; Vissel et al. 1989). Les répétitions de différents chromosomes sont capables de s'agréger à l'interphase pour former une structure compacte nommée chromocentre. Ces répétitions sont fortement méthylées et participent à la structure du chromocentre. En effet, le traitement de cellules eurythroleucémiques murines par la 5-aza entraîne une réexpression des satellites mineurs et une décompaction des chromocentres (Bouzinba-Segard et al. 2006). La méthylation de l'ADN au niveau de ces

éléments répétés semble être conduite par l'environnement chromatinien. La délétion de *Suv39h1* et *Suv39h2*, deux histones méthyltransférases, dans des cellules ES de souris provoque une diminution du niveau de méthylation de l'ADN et elle est associée à une activation de la transcription des répétitions (Lehnertz et al. 2003).

Par ailleurs, chez l'Homme, le syndrome ICF (Immunodeficiency, Centromeric instability, Facial anomalies) est une maladie génétique touchant le gène *Dnmt3b* et qui est caractérisée par une hypométhylation des répétitions péricentriques associée à une instabilité chromosomique (Elrich et al. 2008).

La méthylation de l'ADN joue également un rôle dans la formation du nucléole. Les cellules *Dnmt1*<sup>-/-</sup> présentent, en effet, une désorganisation du nucléole (Espada et al. 2007). Or le nucléole est constitué de séquences répétées codant pour les précurseurs des ARNs ribosomaux et agrégées dans ce compartiment nucléaire spécialisé. Il a été montré qu'un ARN non codant interagit avec le promoteur de l'ADN ribosomal permettant ainsi le recrutement de DNMT3b et l'inhibition de la transcription des ARNs ribosomaux (Schmitz et al. 2010).



**Figure 13. Structure des centromères chez la souris (Guenatri et al. 2004).** Les centromères sont constitués d'éléments répétés appelés satellites mineurs et satellites majeurs.

### 3.5 Rôle de la méthylation de l'ADN dans le corps des gènes

Les profils de méthylation à l'échelle du génome ont permis de révéler qu'une quantité importante de méthylation de l'ADN est en fait localisée dans le corps des gènes (Lister et al. 2009 ; Laurent et al. 2010). Cette distribution est conservée à travers l'évolution et reflète probablement une fonction ancestrale de la méthylation de l'ADN (Feng et al. 2010 ; Zemach et al. 2010). La méthylation dans le corps des gènes est fréquemment retrouvée dans les gènes exprimés de façon ubiquitaire et elle est corrélée à la transcription (Ball et al. 2009). Bien que surprenantes, ces données sont cohérentes avec l'enrichissement de la marque H3K9me3 dans le corps des gènes (Vakoc et al. 2005). Une explication possible serait que la méthylation dans le corps des gènes régule l'efficacité de la machinerie d'élongation et empêche l'initiation de la transcription de promoteur interne alternatif (Maunakea et al. 2010).

Une autre fonction hypothétique de cette méthylation serait de réguler l'épissage via le recrutement du spliceosome par des MBPs. Par ailleurs, la méthylation de l'ADN pourrait permettre la modification des histones et altérer la position des nucléosomes. On sait d'ores et déjà que les marques des histones peuvent influencer l'épissage (Luco et al. 2010). Par ailleurs, des travaux supportent ces hypothèses en montrant un défaut d'épissage dans les cerveaux de souris mutantes pour MeCP2 (Young et al. 2005)

### 3.6 L'empreinte parentale

En 1984, deux équipes d'embryologistes de la souris révèlent un résultat surprenant. Leurs expériences de transplantations de noyaux dans des œufs fécondés montrent que le génome paternel et le génome maternel ne sont pas interchangeables : un œuf ne peut se développer à terme que s'il contient une copie du génome paternel et une copie du génome maternel, bien que ces deux génomes soient strictement identique du point de vue de leur séquence. En effet, chez les embryons gynogénètes le développement du placenta est perturbé tandis que les androgénètes présentent des défauts de développement de l'embryon (McGrath et Solter 1983 ; Surani et al. 1984).

Cette découverte montre qu'il existe une asymétrie parentale et que les génomes parentaux ne sont pas fonctionnellement équivalents. Le terme d'empreinte parentale est alors suggéré. Les expériences qui suivirent permirent d'identifier les premiers gènes soumis à l'empreinte et démontrèrent que ce phénomène correspond à l'expression monoallélique d'un gène dépendant de l'origine parentale.

Il s'est avéré par la suite que la méthylation de l'ADN est un mécanisme moléculaire essentiel à la régulation de l'empreinte en déposant cette marque sur une région régulatrice, de façon différentielle entre le chromosome paternel et le chromosome maternel. La lecture de cette empreinte détermine l'expression d'un allèle, et la répression de l'autre. De nombreux efforts ont été effectués afin d'identifier des régions différentiellement méthylées entre les allèles et dix ans après la découverte des premiers gènes *Igf2* et *Igf2r*, exprimés respectivement paternellement et maternellement, on compte aujourd'hui plus de 100 gènes soumis à l'empreinte chez la souris ([www.mousebook.org](http://www.mousebook.org)) et presque autant chez l'Homme (Morison et al. 2005).

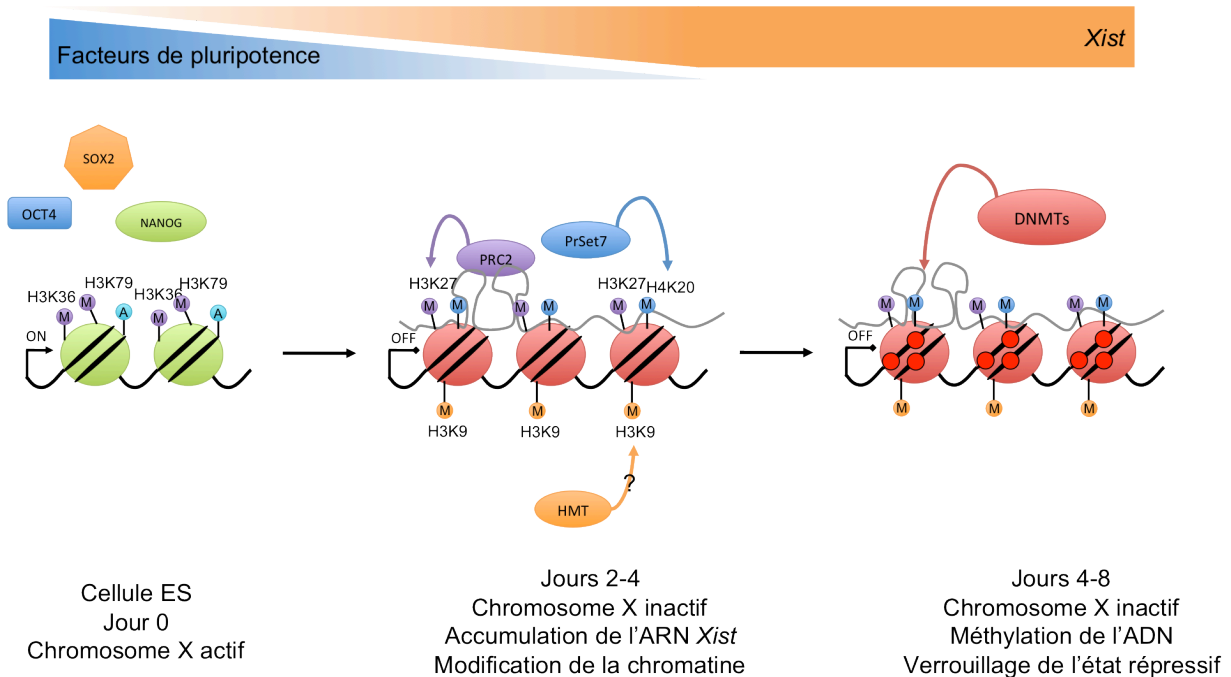
La régulation de l'expression monoallélique de ces gènes est essentielle pour le développement, la croissance fœtale, le métabolisme nutritif et le comportement chez l'adulte. Chez l'homme, de nombreuses maladies sont associées à défauts de régulation de l'empreinte tel que le syndrome de Prader-Willi, le syndrome d'Angelman et le syndrome de Silver-Russell.

La plupart des gènes soumis à l'empreinte sont regroupés dans des régions de 100 à 1000 kb. Ces régions sont riches en CpGs et sont différentiellement méthylées entre les deux chromosomes parentaux et permettent de réguler un ensemble de gènes soumis à l'empreinte, on parle alors de régions contrôlant l'empreinte (ICR : Imprinting Control Region). Il a en effet été montré que la délétion de *Dnmt1* provoque la réexpression biallélique des gènes soumis à l'empreinte dans les embryons de souris (Li et al. 1993). Toutefois, d'autres facteurs et modifications épigénétiques

entrent en jeu dans les mécanismes d’empreinte comme les modifications des histones, les ARNs non codants et également le facteur CTCF.

Les mécanismes de l’empreinte nécessitent trois étapes principales : l’établissement, la maintenance et la suppression. Premièrement, l’établissement de la méthylation de l’ADN se produit dans les cellules germinales au cours de la gamétogenèse par les ADN méthyltransférases *de novo* DNMT3A et DNMT3B en association avec DNMT3L. En effet, l’absence de DNMT3L dans les oocytes cause une hypométhylation de l’ICR des gènes soumis à l’empreinte maternelle tels que *SNRPN*, *Airn* et *PEG3* (Bourc’his et al. 2001 ; Kaneda et al. 2010). Ce défaut de méthylation induit soit une expression biallélique soit une absence d’expression démontrant le rôle critique de DNMT3L dans l’établissement de l’empreinte maternelle. Tandis que l’invalidation chez les mâles provoque un arrêt prématuré de la spermatogenèse et ne permet donc pas l’analyse du statut de méthylation des RDMs. Deuxièmement, les marques mises en place sont transmises dans le zygote et maintenues au cours du développement et chez l’adulte. Pour cela, la méthylation des gènes soumis à l’empreinte échappe au processus de déméthylation qui a lieu dans l’embryon pré-implantatoire (Reik et al. 2001 ; Sasaki et al. 2008). Ensuite, les gènes non méthylés soumis à l’empreinte doivent faire face à la vague de méthylation *de novo* qui se produit après l’implantation. Bien que les mécanismes ne soient pas entièrement compris, ces gènes semblent être protégés de la méthylation par les modifications des histones H3K4me et l’acétylation (Fournier et al. 2002). Enfin, la méthylation des gènes soumis à l’empreinte doit être effacée dans les cellules germinales primordiales (PGCs) afin de rétablir par la suite le caractère spécifique parental. La déméthylation active semble être le mécanisme privilégié via la déamination de la méthylcytosine en thymidine grâce à la déaminase AID. Les processus de réparation de l’ADN sont alors nécessaires afin de restaurer une cytosine (Popp et al. 2010). Un autre mécanisme possible impliquerait la déméthylation via la conversion de la méthylcytosine en hydroxyméthylcytosine via les protéines de la famille TET (Wu et al. 2010).





### Figure 14. Dynamique de l'inactivation du chromosome X.

Les cellules ES au stade non-différencié possèdent deux chromosomes X actifs. La chromatine est dans un état permissif pour la transcription avec la méthylation de H3K36, H3K79 et l'acétylation des histones. En revanche, le recrutement des facteurs de pluripotence inhibe la transcription de l'ARN *Xist*. A partir du jour 2, *Xist* est exprimé et s'accumule autour d'un des chromosomes X. Un changement d'état chromatinien s'effectue, les marques répressives sont mises en place via notamment PRC2. L'état répressif est alors verrouillé entre le jour 4 et 8 par la mise en place de la méthylation de l'ADN qui assure le maintien de l'inactivation du chromosome X.

### 3.7 Inactivation du chromosome X

Tout comme l’empreinte parentale, l’inactivation du chromosome X correspond à un mécanisme d’expression monoallélique. Elle apparaît comme indispensable, permettant la compensation de dose entre mâle et femelle chez les mammifères.

L’inactivation du chromosome X est un exemple particulièrement intéressant pour notre compréhension de la régulation de l’expression des gènes (figure 14). En effet, le processus d’inhibition fait appel aussi bien à la machinerie « classique » de régulation via le recrutement de facteurs de transcription, qu’aux différents mécanismes épigénétiques. Dans un premier temps, des facteurs de pluripotence comme Nanog, Oct4, Rex1 et Sox2 régulent la transcription d’un long ARN non codant, *Xist*. Au cours de la différenciation des cellules ES, ces facteurs sont inhibés ce qui permet l’expression de *Xist* qui joue alors un rôle essentiel dans l’initiation de l’inactivation par sa capacité à s’accumuler sur l’un des deux chromosomes X (Brown et al. 1991 ; Penny et al. 1996). Cette accumulation *in cis* permet la formation d’un compartiment transcriptionnellement inactif dans lequel la machinerie de transcription est absente (Chaumeil et al. 2006). Rapidement après l’accumulation de *Xist*, un changement des modifications des histones est observé. Les marques actives, telles que la méthylation de H3K4 et H3K36 ou l’acétylation des histones H3 et H4, sont perdues tandis que les marques répressives, comme H3K27me3, H3K9me2 et H4K20me, sont mises en place notamment via le recrutement de PRC2 (Plath et al. 2003 ; Silva et al. 2003). Finalement, la méthylation de l’ADN est établie tardivement et permet la maintenance de l’état inactif. En effet, les cellules ES mutées pour *Dnmt1* ou traitées à la 5aza présentent une réactivation partielle du chromosome X inactif (Csankovszki et al. 2001). Par ailleurs, les patientes atteintes du syndrome ICF montrent une instabilité dans la répression du chromosome X inactif (Hansen et al. 2000).

### 3.8 Méthylation de l’ADN et maladies humaines

Comme nous avons pu le voir, la méthylation de l’ADN est un mécanisme crucial dans la régulation de nombreux processus biologiques. Son implication dans le développement embryonnaire, la différenciation,

l'inactivation du chromosome X, l'empreinte parentale et le maintien de l'intégrité du génome soulève alors la question du rôle de cette modification épigénétique dans les maladies humaines et dans les cancers.

### 3.8.1 Défaits de méthylation de l'ADN

Parmi les différentes maladies liées à la méthylation de l'ADN, un grand nombre correspond à des défauts d'empreinte. C'est en effet le cas pour le syndrome Beckwith-Wiedemann caractérisé notamment par une croissance excessive et une prédisposition à développer des tumeurs embryonnaires telles que les tumeurs de Wilm (Elliott et al. 1994). Le locus responsable de la maladie se trouve sur le chromosome 11 et s'étend sur près de 1 Mb incluant plusieurs gènes soumis à l'empreinte. Cette région contient deux ICRs contrôlant les loci *IGF2/H19* et *KCNQ1*. La perte d'empreinte causée par l'absence de méthylation sur l'ICR maternel de *KCNQ1* et l'hyperméthylation de l'allèle maternel sur le locus *IGF2/H19* sont responsables, respectivement, des malformations anatomiques et de la tumorigénèse (Moulton et al. 1994 ; Steenman et al. 1994).

Le syndrome de Prader-Willi est autre exemple de défaut d'empreinte. Les individus atteints présentent une obésité morbide, ainsi que des retards mentaux et des problèmes de comportement (Nicholls et al. 2001). Bien que la majorité des défauts moléculaires proviennent d'une délétion du chromosome 15 paternel et d'une disomie uniparentale, il a également été observé une perte d'empreinte associée à une méthylation *de novo* de l'allèle paternel (Tsai et al. 1999 ; Butler et al. 2002).

Le syndrome d'Angelman touche la même région soumise à l'empreinte que le syndrome de Prader-Willi et il est caractérisé par un retard mental, un déficit de la parole et un comportement anormal (Nicholls et al. 2001). Il est causé par une perte d'expression du gène *UBE3A* exprimé maternellement de façon monoallélique dans le cerveau uniquement et code pour une E3 ubiquitine ligase. Les défauts sont essentiellement génétiques mais dans près de 5% des cas, une perte de la méthylation de l'allèle maternel est en cause (Lossie et al. 2001).

Outre les syndromes liés à l’empreinte parentale, l’altération de la méthylation de l’ADN est également mise en cause dans une maladie à expansion de trinuécléotide : le syndrome de l’X fragile. Comme son nom l’indique, ce syndrome est un défaut lié au chromosome X et se caractérise par un retard mental (Crawford et al. 2001). La région 5’ du gène *FMR1* présente 6 à 52 copies des répétitions CGG chez un individu sain. Or chez les patients, le nombre de répétitions dépasse les 200 et les séquences deviennent méthylées induisant alors une répression du gène *FMR1* (Oberle et al. 1991).

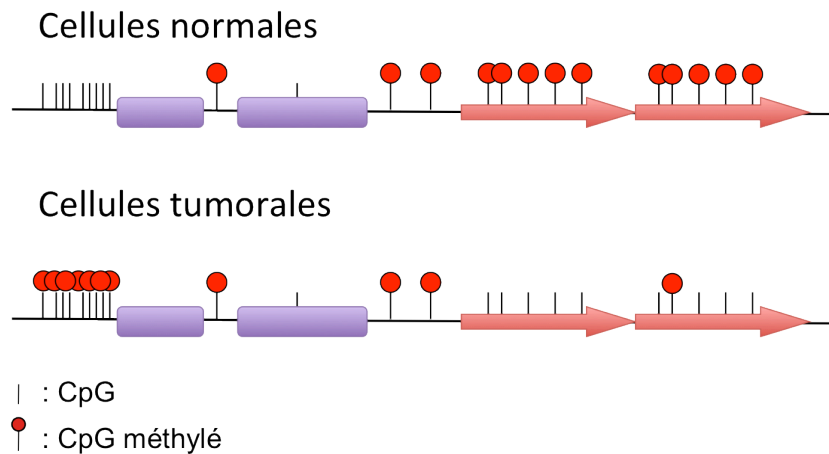
A l’inverse, la dystrophie facio-scapulo-humérale se caractérise par une diminution du nombre de répétition. En effet, un individu sain présente entre 11 et 150 copies de l’élément répété D4Z4 tandis que les personnes atteintes ne possèdent qu’une dizaine de copies. Par ailleurs, la diminution du nombre de répétitions est associée une hypométhylation des éléments D4Z4 (van Overveld et al. 2003). Il a été suggéré que cette hypométhylation soit la cause de la contraction des répétitions et qu’elle serait responsable de la dérégulation des gènes flanquants (Jiang et al. 2003).

### 3.8.2 Défauts des acteurs de la méthylation de l’ADN

Certaines maladies humaines sont également connues pour avoir un défaut des DNMTs. C’est le cas pour le syndrome ICF (Immunodeficiency, Centromeric instability and Facial anomalies) qui est une maladie autosomique extrêmement rare caractérisée par une déficience en immunoglobuline et par un dimorphisme facial (Smeets et al. 1994). D’un point de vue moléculaire, les patients montrent une hypométhylation des répétitions péri-centriques sur les chromosomes 1, 9 et 16. Ce défaut de méthylation provoque des anomalies cytogénétiques dans lesquelles les régions centromériques sont capables de fusionner entre elles. L’origine de cette hypométhylation vient du fait que le gène *DNMT3B* est muté notamment dans sa région catalytique (Ehrlich et al. 2003). Cependant, certains patients présentant une hypométhylation des satellites ne possèdent pas de mutation dans *DNMT3B*, ce qui suggère que d’autres acteurs de la machinerie de méthylation de l’ADN peuvent jouer un rôle. Notamment, des mutations dans le gène codant pour ZBTB24 provoquent

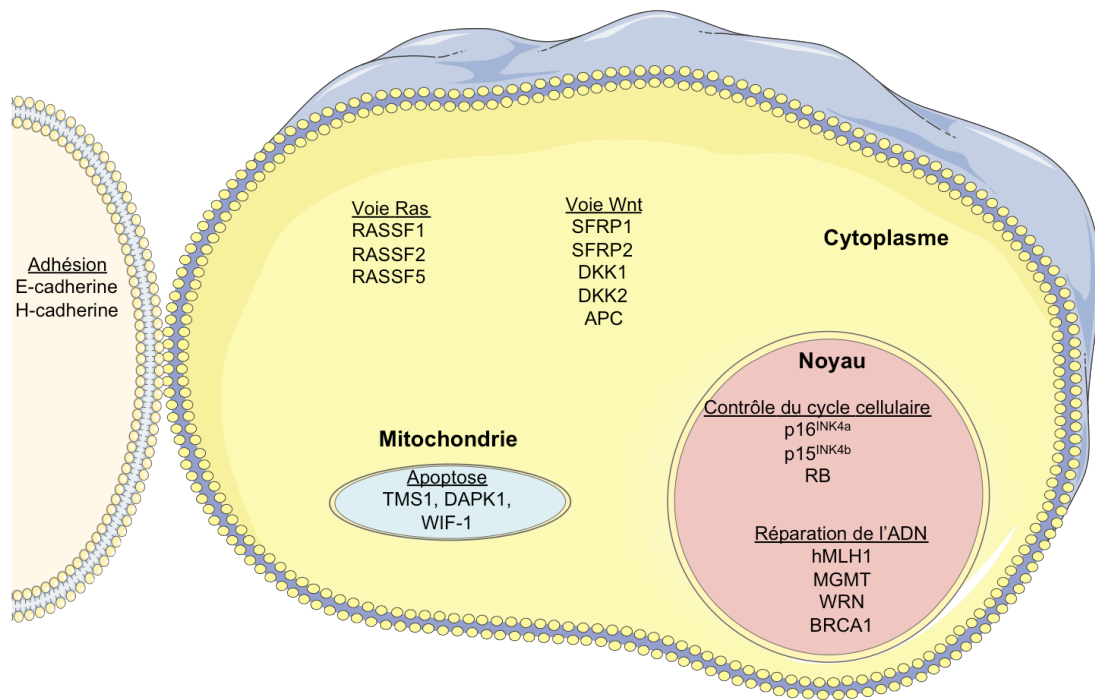
un syndrome ICF associé à une hypométhylation des satellites (de Greef et al. 2011).

Enfin, les mutations des protéines se liant à l'ADN méthylé peuvent provoquer des maladies. Par exemple, le syndrome de Rett est une maladie neuro-développementale qui touche presque exclusivement les femmes. Bien que les causes moléculaires ne sont encore pas très bien comprises, il s'avère que le syndrome de Rett est causé par la mutation du gène *MECP2*, localisé sur le chromosome X (voir Partie I, chapitre 2.2.1)



**Figure 15. Méthylation et cancer.**

Les cellules tumorales présentent une hyperméthylation de nombreux promoteurs dont ceux de gènes suppresseurs de tumeurs, et une hypométhylation du reste du génome. Cette hypométhylation touche particulièrement les séquences répétées. La réexpression qui en résulte contribue à augmenter l'instabilité génomique.



**Figure 16. Méthylation et cancer**

L'hyperméthylation de l'ADN affecte une grande variété de processus biologique comme la régulation du cycle cellulaire, l'apoptose, l'angiogenèse, la réparation de l'ADN, la voie de signalisation Ras.

### 3.8.3 Epigénétique et cancer

En plus de rôles ses physiologiques, la méthylation est liée au cancer. En effet, des patrons de méthylation anormaux sont fréquemment observés de façon précoce lors de la transformation, ce qui suggère une implication causale. Les cellules cancéreuses présentent un niveau total de méthylation réduit par rapport aux cellules saines (Goelz et al. 1985). Mais cette réduction globale recouvre en fait deux phénomènes distincts et inverses. En effet, de nombreux îlots CpG (CGIs) correspondant aux promoteurs de gènes sont en fait hyperméthylés dans les cellules cancéreuses (Ribieras et al. 1994). Ce phénomène touche de nombreux gènes suppresseurs de tumeurs, et inhibe leur expression. Les séquences répétées, au contraire, connaissent une déméthylation dans les cellules cancéreuses (figure 15). Ceci amène leur réactivation, augmente l'instabilité chromosomique, provoque des événements de translocations et de disruption de gènes (Gaudet et al. 2003). C'est notamment le cas pour le rétrotransposon LINE1 qui se retrouve déméthylé dans de nombreux cancers dont des tumeurs du sein, du foie, des poumons et de la vessie (Dante et al. 1992 ; Perrin et al. 2007). Cependant, certains gènes spécifiques deviennent également

hypométhylés dans les cancers. C'est notamment le cas pour le gène *p52*, dont l'expression est contrôlée par les oestrogènes, qui devient hypométhylé au niveau de son enhancer dans de nombreux cancers (Martin et al. 1995 ; Martin et al. 1997 ; Chatagnon et al. 2010).

L'hypométhylation peut également avoir lieu sur les promoteurs de gènes soumis à l'empreinte et induire une activation aberrante. Par exemple, l'hypométhylation du gène *IGF2* (Insulin-like Growth Factor 2) provoque une perte d'empreinte et est associé avec des cancers du sein, du foie, des poumons ou encore du côlon (Ito et al. 2008).

A l'inverse de la perte de méthylation globale, l'hyperméthylation est observée au niveau de CGIs spécifiques. Cette altération de la méthylation de l'ADN affecte une grande variété de processus biologiques comme la régulation du cycle cellulaire, l'apoptose, l'angiogenèse, la réparation de l'ADN, la voie de signalisation Ras ou encore la réponse immunitaire (Esteller 2008). Cependant, les données générées sur l'ensemble du génome tendent à montrer que la méthylation aberrante dans les cancers a lieu en amont, dans les CGIs « shores » (Irizarry et al. 2009 ; Doi et al. 2009).

D'autre part, les tumeurs sont également caractérisées par une hyperméthylation des miARNs qui conduit à une absence d'expression (Saito et al. 2006). L'un des exemples les plus étudié est celui du *miR-124a*, réprimé dans de nombreux cancers et qui contrôle l'activation de *CDK6* (*Cyclin-dependent kinase 6*) et la phosphorylation de *Rb* (Lujambio et al. 2007).

Bien que les causes des défauts de méthylation de l'ADN dans les cancers ne soient pas encore comprises, il semble que la dérégulation des ADN méthyltransférases est un paramètre important dans le développement des tumeurs. En effet, la surexpression de DNMT1 et DNMT3B est observée dans différents cancers. De plus, l'expression des DNMTs peut être régulée par des miARNs comme le *miR-29* qui cible et inhibe directement DNMT3A et DNMT3B et indirectement DNMT1 (Garzon et al. 2009). Par ailleurs, les MBPs semblent également jouer un rôle important dans la tumorigenèse. MeCP2 et MBD2 sont notamment

fortement exprimées dans les cancers du sein et apparaissent donc comme des facteurs clefs à la formation des tumeurs (Billard et al. 2002).

### 3.9 Hydroxyméthylation

La preuve formelle de l'existence de l'hydroxyméthylation (5hmC) chez les mammifères a été apportée en 2009 par des expériences de spectrométrie de masse et de chromatographie (Kriaucionis et al. 2009 ; figure 3). Cette nouvelle modification de l'ADN fait depuis l'objet d'intenses recherches afin de comprendre sa fonction. Bien que la 5hmC soit probablement présente dans tous les types cellulaires, elle apparaît plus abondante dans les neurones et les cellules souches embryonnaires (Kriaucionis et al. 2009 ; Ficz et al. 2011). La 5hmC est également présente à un taux plus faible que la méthylation de l'ADN et les niveaux d'hydroxyméthylation varient selon le type cellulaire de 0,1% à 0,7% du nombre total de cytosine (Globisch et al. 2010).

L'établissement de la 5hmC peut être généré par l'oxydation du groupe méthyle de la méthylcytosine. En effet, la délétion des DNMTs conduit à une réduction de la quantité de 5hmC (Szwagierczak et al. 2010) suggérant que la méthylcytosine est un substrat essentiel pour aboutir à la 5hmC. De la même façon, l'absence de UHRF1 dans les cellules ES provoque une diminution du niveau de 5hmC (Ficz et al. 2011). Afin d'identifier les enzymes responsables de cette oxydation, des recherches d'homologies avec des enzymes appartenant à la famille des hydroxylases ont été effectuées et ont permis d'identifier les protéines de la famille TET (Iyer et al. 2009). La surexpression de TET1 induit une augmentation de la 5hmC et une diminution de la méthylcytosine (Tahiliani et al. 2009). Alors que la délétion de *Tet1* par ARN interférence provoque une réduction de la 5hmC d'environ 40% dans les cellules ES (Wu et al. 2011).

La 5hmC est enrichie avec l'euchromatine au niveau des promoteurs et du corps des gènes. Bien que l'hydroxyméthylation des promoteurs ne soit pas associée avec la transcription, il existe une forte corrélation entre la 5hmC du corps des gènes et leur expression. De façon intéressante, les localisations de la 5hmC et de la méthylcytosine sont exclusives au niveau



des promoteurs alors qu'il existe une certaine redondance dans le corps des gènes (Jin et al. 2011 ; Williams et al. 2011).

Comme pour la méthylation de l'ADN, il a été proposé que l'hydroxyméthylation pourrait jouer un rôle dans la régulation de l'expression des gènes. Cependant, l'hydroxyméthylation ne permet pas de prédire le caractère actif ou inactif de la transcription. D'une part, la 5hmC peut être associée aux marques actives de la transcription dans les régions hypométhylées. En effet, l'anticorrélation entre la 5hmC et la méthylcytosine au niveau des promoteurs suggère que l'hydroxyméthylation peut soulever l'inhibition de la méthylation de l'ADN. D'autre part, la 5hmC est également enrichie sur des régions transcriptionnellement inactives. Il s'avère d'ailleurs que la protéine Tet1 interagit avec le complexe inhibiteur Sin3A et permettrait son recrutement sur certains loci spécifiques (Williams et al. 2011).

La question sur la reconnaissance de la 5hmC par des protéines spécialisées a rapidement été soulevée car leur découverte permettrait une meilleure compréhension sur le rôle biologique de la 5hmC. Les protéines de la famille MBD, MBD1, MBD2, MBD4 et MeCP2, ne semblent pas lier la 5hmC (Jin et al. 2010). En revanche, il a récemment été montré que UHRF1 est capable de reconnaître la 5hmC via son domaine SRA (Frauer et al. 2011). Il sera donc intéressant de savoir si UHRF2 possède également cette propriété ainsi que les protéines à doigts de zinc.

Enfin, il a été proposé que l'hydroxyméthylation soit un intermédiaire dans le processus de déméthylation de l'ADN. Tet1 peut convertir la méthylcytosine en 5hmC qui s'avère être un substrat pour la voie BER de la réparation de l'ADN et conduire à la déméthylation (Guo et al. 2011).





## Chapitre 4 : rôle des protéines liant l'ADN méthylé

### 4.1 Un dogme à réévaluer

Comme nous avons pu le voir, le dogme actuel stipule que les MBPs agissent comme des répresseurs transcriptionnels. Cependant, ce modèle a besoin d'être réévalué sur la base de deux observations :

- Le faible nombre de cibles des MBPs.
- La nature non essentielle de la plupart des MBPs connues.

Cela fait maintenant près de 20 ans que la première MBP, MeCP2, fut identifiée. Malgré les efforts effectués afin d'identifier les gènes qui sont régulés par la méthylation de l'ADN dans les cellules normales, il n'existe qu'un faible nombre de cibles connues. Parmi les exemples les plus clairs, on retrouve MBD1 qui régule le promoteur méthylé de FGF2 dans les cellules souches neurales (Li 2008) et MBD2 qui inhibe le promoteur de  $\rho$ -globine chez le poulet (Kransdorf 2006).

D'autre part, alors que la méthylation de l'ADN est essentielle pour le développement, la plupart des MBPs ne le sont pas : MBD1, MBD2, MeCP2, Kaiso et ZBTB4 peuvent être déplétées individuellement ou en combinaison, sans pour autant avoir des conséquences sur le développement embryonnaire de la souris (Caballero 2009). En revanche, l'absence de UHRF1, qui va conduire à une perte de méthylation, va provoquer un arrêt développemental précoce.

### 4.2 Les protéines liant l'ADN méthylé ne sont pas essentielles

#### 4.2.1 Redondance des MBPs

Une explication possible à la difficulté de trouver les cibles des MBPs peut être une redondance fonctionnelle. Si plusieurs MBPs lient et répriment un promoteur donné, l'absence d'une seule de ces protéines n'induit pas nécessairement une réactivation transcriptionnelle. Cette hypothèse peut également expliquer les phénotypes « modérés » des souris mutantes. Il sera nécessaire par la suite de réaliser de multiples déplétions (TKO, QKO) pour résoudre cette question. Cependant, la souris mutante

pour Kaiso/MBD2/MeCP2 a été générée mais les animaux ne présentent pas de phénotype plus important que les simples mutants.

Toutefois, sachant que les MBPs diffèrent par leur structure, il est concevable de penser qu'elles possèdent également une affinité différentes pour leurs cibles et donc ont des sites préférentiels de liaisons différents. UHRF1 semble préférer l'ADN hémiméthylé alors que MeCP2 a une plus grande affinité pour l'ADN méthylé. De plus, MeCP2, MBD1, Kaiso, ZBTB4 et ZBTB38 lient certaines séquences méthylées plutôt que d'autres *in vitro* (Klose 2005, Clouaire 2010, Sasai 2010). Ces caractéristiques vont donc à l'encontre de la redondance des MBPs.

Une approche expérimentale qui pourrait résoudre cette question serait d'obtenir une carte à l'échelle du génome de toutes les MBPs dans le même type cellulaire.

#### 4.2.2 Existence d'autres MBPs

Il existe de nombreux domaines structuraux capables de lier l'ADN de manière séquence spécifique. La 5-méthylcytosine étant présente chez de nombreux organismes, il n'est pas inconcevable de penser que d'autres domaines aient pu évoluer pour lier cette modification. Il existe d'ailleurs une protéine du virus Epstein-Barr, appelé Zebra, capable de lier les promoteurs méthylés. C'est une protéine bZIP contenant un domaine basique suivi d'une Leucine-Zipper (Petosa 2006). D'autres protéines bZIP non virale pourraient donc avoir les mêmes propriétés.

#### 4.3 Cibles des MBPs

Les profils de méthylation de l'ADN à l'échelle du génome mettent en évidence que dans les cellules normales de mammifère, les promoteurs associés aux CGIs (HCPs) ne sont pas méthylés et donc ne représentent pas de site de liaison pour les MBPs. Toutefois, deux exceptions sont connues, les gènes de la lignée germinale et les gènes du trophoctoderme (Weber et al. 2005 ; Illingworth et al. 2007). Les MBPs se présentent comme étant de bons candidats dans le processus de répression de ces gènes, mais cela n'a pas encore été démontré.

Les MBPs pourraient agir sur d'autres régions du génome et notamment sur les CGIs « shores » qui présentent un niveau de méthylation différentiel entre les tissus sains.

Les MBPs pourraient également réguler les promoteurs avec un pourcentage plus faible de CpGs. En particulier les promoteurs avec un contenu « intermédiaire » de CpGs (ICPs) et caractérisés pour avoir des niveaux de méthylation différents entre les tissus. Une hyperméthylation de certains ICPs serait cohérente avec la répression des gènes.

L'hétérochromatine péricentromérique pourrait être une cible pour les MBPs. Chez les mammifères, cette structure est fortement méthylée mais la transcription de cette région permet de contrôler le fonctionnement du centromère via un ARN non codant (Bouzinba-Segard 2006, Ferri 2009). L'absence de MeCP2 dans le cerveau induit la transcription des satellites majeurs, répétitions caractéristiques des centromères (Skene 2010). Par ailleurs, ATRX, un partenaire de MeCP2, stimule la transcription des répétitions péricentromériques en favorisant l'incorporation du variant d'histone H3.3 (Drané 2010). Or ATRX semble nécessaire à la localisation de MeCP2 aux chromocentres au moins dans les neurones (Nan 2007). Une possibilité serait que ATRX régule négativement MeCP2 pour permettre la transcription des satellites majeurs. De plus, l'hétérochromatine péricentromérique attire la plupart des MBPs dans des expériences de transfection. On peut donc supposer que l'absence d'une ou plusieurs MBPs pourrait causer des défauts de régulation de ces régions dans les souris mutantes.

Il semblerait que la méthylation des éléments subtelomériques est nécessaire à la fonction des télomères (Gonzalo 2006). Des MBPs pourraient alors lier ces éléments et participer à l'assemblage de la structure hétérochromatique.

Enfin, les corps des gènes sont fortement méthylés et représentent donc une cible potentielle pour les MBPs.

#### 4.4 MBPs : activateurs transcriptionnels ?

Une analyse du transcriptome des hypothalamus et des cervelets de souris, présentant des niveaux variables de MeCP2, révèle qu'un grand nombre de gènes sont en fait activés par MeCP2 et réprimés en son absence (Ben-Shachar 2009). Des expériences de ChIP montrent qu'il s'agit de cible directe pour MeCP2. Ces observations peuvent être expliquées par la capacité de MeCP2 à recruter un activateur transcriptionnel CREB (Chahrour 2008).

Des tests de rapporteur montrent que MeCP2 et MBD2 peuvent promouvoir l'activation de certaines cibles méthylées (Ego 2005 et Matsumura 2010). Certains virus sont capables de changer la fonction des MBPs de l'hôte en activateurs transcriptionnels via une protéine virale. C'est le cas de Tax pour le virus HTLV-1 (Ego 2005) et de LANA pour le virus de l'herpès associé à un sarcome de Kaposi (Matsumura 2010).

Des tests d'interaction protéines-protéines montrent que MBD2 (Angrisano 2006) et UHRF1 (Achour 2009) interagissent avec des coactivateurs et des HATs.

Bien que le mécanisme soit indirect, MBD4 est requis dans le processus d'activation transcriptionnelle via son rôle dans la déméthylation de l'ADN (Rai 2008).

De nombreux facteurs de transcription comme p53, Myc, YY1 peuvent activer ou inhiber la transcription selon le contexte. Alors la possibilité que les MBPs puissent activer la transcription n'est pas déraisonnable. Si certaines MBPs, comme MeCP2, ont les deux fonctions, le stimulus causant leur changement d'activité n'est pas clair. Des modifications post-traductionnelles, des changements structuraux induits par la liaison à l'ADN ou encore l'action de cofacteurs peuvent être à l'origine de cette bifonctionnalité.







## **PARTIE II : RECHERCHE DE NOUVELLES PROTEINES HUMAINES SE LIANT A L'ADN METHYLE**

### **Chapitre 1 : les différentes approches abordées**

Comme nous avons pu le voir précédemment, trois familles de protéines sont connues pour lier l'ADN méthylé, les protéines de la famille MBD (MBD1, MBD2, MBD4 et MeCP2), les protéines de la famille des Kaiso-like-zinc-finger (Kaiso, ZBTB4 et ZBTB38) et les protéines de la famille SRA (UHRF1 et UHRF2). A l'exception d'UHRF1, aucune de ces protéines n'est essentielle. Par opposition, la méthylation elle-même est essentielle : chez la souris, la perte des ADN-méthyltransférases (DNMTs), est létale et chez l'Homme seules les mutations atténuantes semblent viables (elles sont cependant à l'origine du syndrome ICF). Il est donc légitime de penser que d'autres protéines sont capables d'interpréter le signal de méthylation de l'ADN.

La découverte, au sein du laboratoire, de ZBTB4 et ZBTB38 (Filion et al.) valide clairement l'idée que des protéines liant l'ADN méthylé restent à identifier, et qu'elles jouent des rôles importants chez les mammifères (Weber et al. 2008).

Notre objectif à terme est de mieux comprendre les interprètes de la méthylation de l'ADN, un processus épigénétique fondamental pour l'activité du génome. Pour cela, nous souhaitons élargir et systématiser notre recherche de protéines humaines liant l'ADN méthylé, puis caractériser les protéines obtenues. Dans cette optique, nous avons opté pour deux types d'approches distinctes, une approche génétique et une approche basée sur la littérature.

#### **1.1 Approche génétique**

Nous proposons une nouvelle stratégie reposant sur le "phage-display" : une banque de phages exprimant à leur surface des protéines de fusion codées par des cDNA humains est mise en contact avec des billes

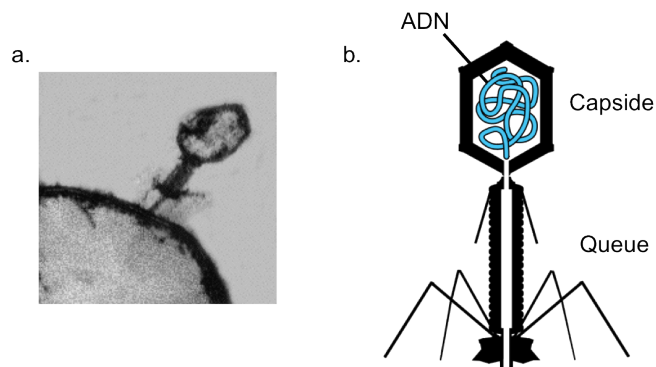
magnétiques recouvertes d'ADN méthylé. De l'ADN non méthylé présent en excès permet d'éliminer les interactions non-spécifiques. Cette approche s'inspire de techniques anciennes qui ont fait leurs preuves (Singh et al. 1988), et a été récemment appliquée pour identifier des facteurs de transcription se liant à une séquence consensus (Cicchini et al. 2002). Elle est rapide et flexible : les séquences méthylées peuvent être variées facilement, de même que les conditions physico- chimiques d'incubation ADN-protéine.

## 1.2 Approche basée sur la littérature

Nous tirerons pleinement parti des résultats issus d'approches pan-génomiques. Ainsi, un article récent a identifié les protéines nécessaires à la méthylation anormale du promoteur Fas suite à la transformation de cellules par H-ras (Gazin et al. 2007). Des expériences d'immunoprécipitation de la chromatine (ChIP) montrent que certaines de ces protéines se lient spécifiquement au promoteur méthylé mais pas avec le promoteur non méthylé. Ce sont donc d'excellents candidats pour être des protéines reconnaissant l'ADN méthylé.

La méthylation de l'ADN contrôle de façon cruciale l'activité du génome dans les cellules normales, et sa perturbation contribue à la transformation tumorale. En étudiant les protéines qui se lient à l'ADN méthylé chez l'Homme, nous espérons, à terme, mieux comprendre les étapes précoces de la transformation pour pouvoir les détecter et les contrecarrer.





**Figure 17. Structure du phage T7.**

**A :** Phage T7 infectant une bactérie (<http://bio.research.ucsc.edu>). **B :** Le phage T7 est composé d'une capside renfermant un génome d'ADN double brin d'environ 40 Kpb. Il possède une queue permettant la liaison avec les bactéries et l'injection de son ADN à la manière d'une seringue.

## Chapitre 2 : approche génétique, le phage display

### 2.1 Introduction

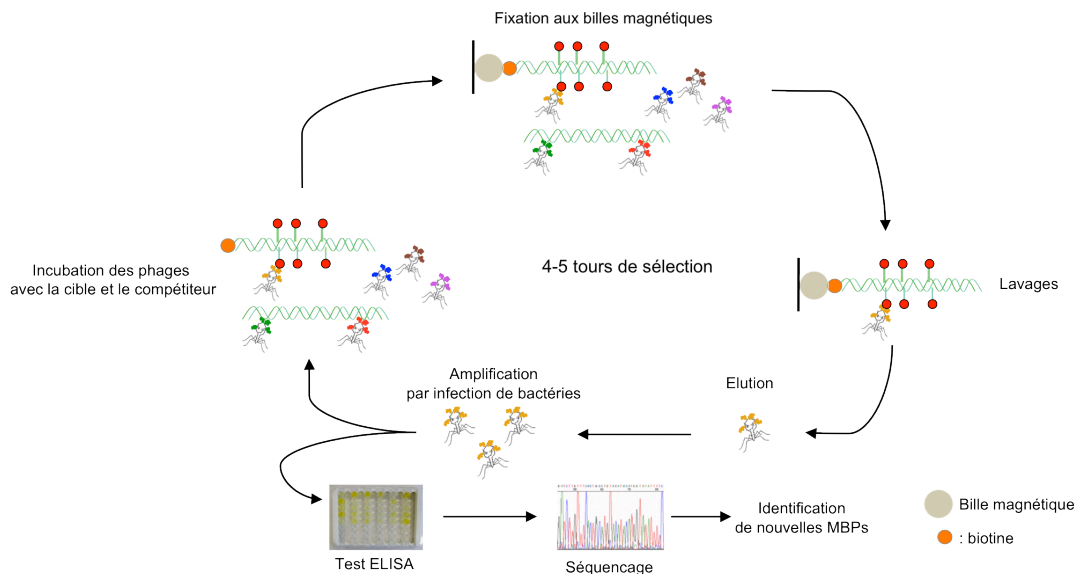
Le phage display est une technique de sélection *in vitro* extrêmement puissante dans l'étude des interactions protéines-protéines, protéines-peptides, protéines-ADN ou protéines-ARN. Le phage display repose sur la fusion d'un gène d'intérêt avec un gène codant une protéine exprimée à la surface du bactériophage (ou phage ; figure 17), la capside, permettant ainsi l'expression de peptides (ou protéines) d'intérêt à la surface de celui-ci.

Comme pour le double hybride, des banques de phages peuvent ainsi servir à réaliser des cribles dans le but d'identifier des protéines (ou peptides) d'intérêt ayant une affinité spécifique pour une protéine, un ADN ou un ARN *in vitro*. Les clones sélectionnés peuvent alors être analysés et la nature du gène identifié par séquençage. Depuis son invention par George Smith en 1985 (Smith et al. 1985), la technologie a beaucoup évolué et permet d'identifier des polypeptides avec de nouvelles propriétés. Cette technologie a notamment démontré son intérêt par la possibilité de sélectionner des anticorps humains et donc d'être une alternative aux techniques classiques au clonage des anticorps (Barbas et al. 1991). Les industries pharmaceutiques et biotechnologiques ont alors repris le phage display dans le but de découvrir des anticorps, des protéines recombinantes ou encore d'optimiser des molécules existantes. Pour exemple, HUMIRA est le premier anticorps monoclonal capable de réduire les symptômes de l'arthrite rhumatoïde en ciblant le facteur TNF- $\alpha$ , découvert par la technologie du phage display et dont les ventes ont dépassées les 2 milliards de dollars en 2006 (Lawrence et al. 2007).

Les bactériophages furent identifiés il y a presque un siècle par Félix d'Hérelle à l'Institut Pasteur comme étant des virus capables d'infecter des bactéries (F. d'Hérelle, 1917). Les phages possèdent une enveloppe protéique appelée capside renfermant le matériel génétique qui peut être un ARN ou ADN simple ou double brin (figure 10b). Chaque phage a la propriété d'infecter un certain type de bactéries. Par la suite, nous nous

intéresserons au phage T7 qui est un virus à ADN double brin infectant spécifiquement *Escherichia coli*. La grande flexibilité et la possibilité de modifier aisément le génome des phages en font un outil génétique extrêmement puissant. En effet, il est possible de remplacer de grandes parties de leur génome sans pour autant changer la viabilité des phages (Chan et al. 2005). Ceci peut être expliqué par le fait que leur survie ne dépend que d'un ensemble restreint de gènes nécessaires à leur réplication et à l'encapsidation. Les bactériophages peuvent avoir un cycle lytique ou un cycle lysogénique ou les deux.

Les phages lytiques, comme le phage T7, ont un cycle divisé en plusieurs étapes comprenant l'adsorption, la pénétration de l'ADN, la duplication et l'encapsidation. Pendant l'adsorption, le phage T7 s'attache à la bactérie via la reconnaissance de lipopolysaccharides (LPS) présents à la surface de la cellule hôte. Ensuite, le génome du phage est injecté dans la bactérie et doit échapper aux mécanismes de défense de la cellule hôte. L'ADN phagique est alors rapidement répliqué en utilisant essentiellement les ressources en nucléotides de l'hôte dont il a précédemment dégradé le génome. L'ADN est alors encapsidé et le phage formé suite à la synthèse des protéines de structure par la machinerie de l'hôte. Les phages sont finalement libérés lorsque la bactérie est dégradée par des lysosymes d'origine phagique ciblant les peptidoglycanes de la membrane cellulaire.



**Figure 18. Principe du crible par phage display.**

Les phages sont incubés avec une sonde méthylée et biotinylée ainsi qu'avec un compétiteur non méthylé. Les phages sont alors sélectionnés par fixation sur une matrice recouverte de Streptavidine (ici, une bille magnétique). Les phages sélectionnés sont ensuite amplifiés avant de réaliser un nouveau tour de sélection. Après 4 à 5 tours de sélection, les phages sont analysés par un test ELISA afin d'évaluer leur affinité pour la sonde méthylée. Les clones phagiques présentant la plus forte interactions avec l'ADN méthylé sont alors séquencés.

## 2.2 Principe du crible par phage display

Le crible par phage display peut être divisé en trois grandes étapes : la création ou l'obtention d'une banque d'ADNc exprimant des polypeptides à la surface des phages, la sélection, et enfin l'analyse des clones obtenus. Nous reviendrons par la suite sur l'étape d'obtention d'une banque de phages (figure 18).

Le phage display repose sur la sélection de clones spécifiques à travers l'incubation d'une banque de phages avec leur cible (dans notre cas, il s'agit d'un ADN méthylé) afin de créer des liaisons *in vitro*. Les phages interagissant avec la cible sont alors soumis à des lavages et les phages retenus sont élués et amplifiés par infection de bactéries. Le lysat correspondant est ensuite récupéré et un nouveau tour de sélection est réalisé jusqu'à l'enrichissement d'un nombre restreint de phages. Les phages ainsi sélectionnés peuvent alors être individualisés en infectant des bactéries étalées sur une boîte de Pétri contenant de l'agar. Les phages, en



dégradant les cellules hôtes, forment des plages de lyse qui correspondent à un clone phagique exprimant un ADNc particulier qui sera caractérisé par la suite par ELISA et/ou séquençage (figure 22).

L'avantage du crible par phage display réside dans l'étape de sélection qui s'avère extrêmement flexible puisque l'expérimentateur peut changer de nombreuses conditions d'incubations, de lavages ou encore d'élutions. Ceci est rendu possible par l'extraordinaire stabilité des phages qui peuvent s'accommoder de larges variations de pH, des concentrations élevées de détergents ou encore de dénaturants.

Après sélection des phages spécifiques, il est nécessaire de les caractériser et de les identifier. Une possibilité s'avère tout simplement être le séquençage des clones enrichis. L'autre possibilité est de caractériser les phages par ELISA. Les cibles sont immobilisées dans des plaques de 96 puits et le surnageant de chaque phage sélectionné est ajouté afin de détecter une liaison spécifique à l'aide d'un anticorps anti-phage. En plus de confirmer la spécificité du crible, la technique ELISA peut également déterminer l'affinité de liaison relative entre les clones.

Le phage display se présente donc comme un outil génétique adaptable permettant le clonage et l'obtention rapide d'une banque de phage grâce à la facilité de manipulation de son génome. Il s'avère robuste et hautement sélectif de par la possibilité de changer les conditions de stringence. Enfin, il est rapide grâce aux propriétés d'amplification d'un clone particulier.

## **2.3 Développement du crible par phage display**

### **2.3.1 Choix de l'espèce de bactériophage et de la banque d'ADNc**

Le bactériophage communément utilisé est le phage filamenteux M13 qui se propage par un mécanisme non-lytique. Ceci a pour conséquence que toutes les protéines constituant la capside doivent être exportées dans la bactérie hôte. Donc seules les protéines capables d'être injectées dans la cellule seront présentes à la surface du phage. Or cette propriété dépend de plusieurs paramètres tels que la présence d'un peptide signal, la séquence et le repliement des protéines à présenter sur la capside. Par conséquent, un

nombre non négligeable de protéines codées par une banque d'ADNc ne sera pas capable d'être affiché sur le phage M13. Afin de contourner cette limitation, nous avons décidé d'utiliser le phage T7 lytique dont l'assemblage de la capsid est effectué dans le cytoplasme de la bactérie hôte.

Le phage T7 présente l'avantage d'exprimer une dizaine de copies de polypeptides d'environ 1200 acides aminés. Sachant que les domaines MBD et les doigts de zinc représentent environ 80 acides aminés et que le domaine SRA est long de 150 acides aminés (figure 10), cette propriété est donc suffisante pour répondre à notre question biologique, à savoir l'identification de nouveaux motifs et/ou de nouvelles protéines se liant à l'ADN méthylé.

Par ailleurs, des banques de phages ont été générées à partir d'ADNc de tissus humains et sont désormais commercialisées (T7 Select libraries, *Novagen*). Nous avons donc décidé de tirer profit de cet outil afin de nous consacrer aux étapes de sélection et d'analyse. La banque utilisée au cours des différents cribles correspond à des ADNc de cerveau humain sain (T7 Select libraries, cat. 70637-3, *Novagen*). En effet, le cerveau présente l'avantage d'exprimer fortement *MeCP2* et également *Kaiso* et *Zbtb4* (Skene et al. 2010). La sélection de ces protéines lors du crible serait donc un bon indicateur de l'efficacité de la technique. De plus, chez l'homme et la souris, la méthylation de l'ADN est plus élevée dans le cerveau que dans les autres tissus. Il est donc légitime de penser que d'autres protéines de liaison à l'ADN méthylé y soient exprimées afin d'interpréter le signal de méthylation.

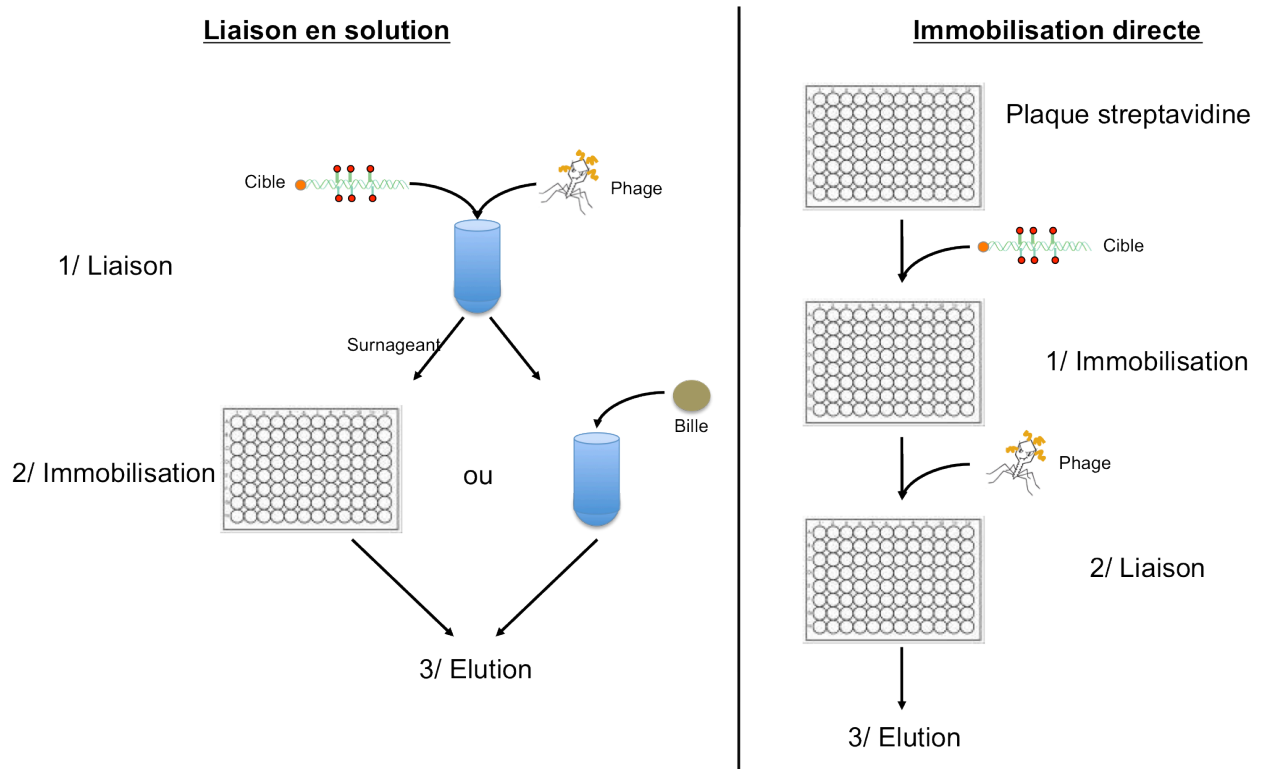
### 2.3.2 Détermination des conditions expérimentales

Bien que le phage display ait fait ses preuves dans la mise en évidence d'interactions protéines-acides nucléiques, aucune étude à ce jour n'a rapporté un crible utilisant cette technique entre des protéines et un ADN méthylé. Il est donc nécessaire, dans un premier temps, de savoir si le phage display est adapté à notre finalité. Si oui, il est par la suite important

d'optimiser le crible afin de minimiser le bruit de fond et les faux positifs. Plusieurs critères rentrent alors en compte :

- Le choix de l'approche. L'interaction ADN méthylé-protéines peut avoir lieu soit en solution soit lorsque l'ADN méthylé est préalablement immobilisé sur la matrice (figure 19).
- Le choix de la matrice qui lie le complexe ADN méthylé-protéines : la matrice étant le support de présentation de la cible, elle doit être choisie en fonction du bruit de fond. L'immobilisation d'une cible biotinylée sur une matrice recouverte de streptavidine semble la solution la plus adéquate en raison de la forte interaction entre les deux molécules ( $K_d \approx 10^{-15}$  M). Principalement, deux types de matrice recouverte de streptavidine peuvent être utilisés, les billes magnétiques ou les plaques 96 puits.

Chaque approche est testée avec les différentes matrices et le bruit de fond est mesuré. Pour cela, nous avons généré un contrôle positif en insérant les acides aminés 1 à 183 de MeCP2, contenant le domaine de liaison à l'ADN méthylé, dans le vecteur T7 Select 10-3b (*Novagen*), fusionné à la protéine de capsid 10B. Une réaction d'encapsidation *in vitro* est effectuée suivie par une amplification en plaque de lyse. Les phages recombinants sont finalement sélectionnés par PCR. Par la suite, nous parlerons de phages positifs pour les phages exprimant le domaine de reconnaissance à l'ADN méthylé de MeCP2. D'autre part, nous avons également sélectionné un phage négatif pour sa liaison à l'ADN à partir de la banque d'ADNc de cerveau humain. L'insert du phage sélectionné est identifié par séquençage et code pour la Thioredoxin domain containing 13 (TXNDC13). Le cadre de lecture dans lequel l'ADNc est inséré provoque l'apparition d'un codon stop précoce conduisant à l'expression d'une protéine tronquée. Par ailleurs, la protéine TXNDC13 n'est pas connue pour interagir avec les acides nucléiques. Ce contrôle s'avère donc parfaitement adapté aussi bien pour le choix de la matrice que pour l'adaptation de cette technique à notre question biologique.



**Figure 19. Les différentes approches par phage display.**

Deux approches distinctes peuvent être envisagées. Une approche dans laquelle les interactions phages-ADN méthylé sont réalisées en solution et une approche où les phages sont incubés avec une cible préalablement immobilisée sur la matrice.

La capacité des phages positifs et négatifs à lier l'ADN, l'ADN méthylé ou la matrice seule est alors quantifiée selon les approches en variant les différents tampons d'incubation et les différents agents bloquants. Pour cela, nous incubons séparément les phages positifs et les phages négatifs avec un ADN méthylé, un ADN non méthylé ou sans ADN. Après le quatrième tour de sélection, les phages interagissant avec leur cible sont élués et le nombre de phages enrichis dans chaque condition est quantifié (ou titré). Afin de titrer le nombre de phages, des bactéries sont infectées par l'éluat et étalées sur des boîtes de Pétri. Le nombre de phages présents dans l'éluat est alors déterminé par le nombre de plages de lyse formées (ou pfu : plaques forming units).

La première observation que nous pouvons faire est que quelque soit l'approche utilisée, l'incubation après immobilisation directe ou en solution, l'effet sur le bruit de fond est négligeable. En revanche, nous avons pu noter que le choix de la matrice a un effet important sur les interactions non spécifiques.

#### Les billes magnétiques recouvertes de streptavidine

Les billes présentent un fort bruit de fond lors de la liaison du phage positif à l'ADN non méthylé et également à la matrice seule comparé à la liaison avec l'ADN méthylé. L'utilisation de différents agents bloquants (BSA, lait, ADN de sperme de saumon) pour saturer les billes ainsi que différents tampons de liaison (PBST, BW, TENN ; voir Table 12 ) n'ont pas permis la réduction du bruit de fond. Par ailleurs, le phage négatif lie tout aussi bien l'ADN qu'il soit méthylé ou non et les billes. Enfin, le phage positif n'est pas enrichi comparé au phage négatif pour une même cible (Table 4).

#### Les plaques streptavidines

Les plaques semblent plus adéquates pour l'enrichissement spécifique du phage positif. Premièrement, le phage positif est plus affiné pour l'ADN méthylé que pour la matrice seule. Deuxièmement, le phage positif a une plus grande affinité que le phage négatif pour la cible méthylée (Table 4).

Sur la base de ces résultats, les plaques permettent le meilleur enrichissement du phage positif tout en réduisant au minimum le bruit de fond. Nous avons donc fait le choix d'utiliser les plaques streptavidines comme matrice dans les expériences suivantes.

**Table 4. Effets des matrices sur la sélection des phages.**

Billes	Positif	Négatif
DNA me <sup>+</sup>	+++	+++
DNA me <sup>-</sup>	+++	+++
No DNA	++	++

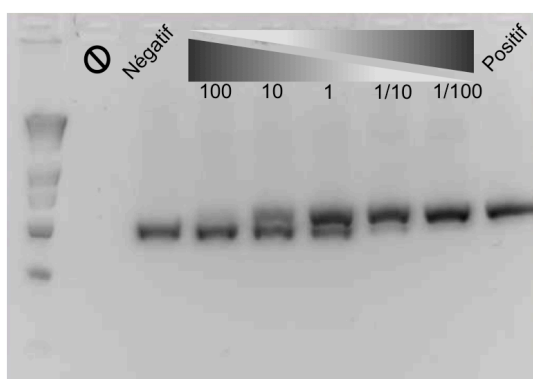
Plaques	Positif	Négatif
DNA me <sup>+</sup>	++++	++
No DNA	++	++

Ces tables résument les résultats d'enrichissement obtenus des phages positifs et négatifs avec les différents tampons (**Table 12**). ++ : 10<sup>4</sup> pfu/ml ; +++ : 10<sup>5</sup> pfu/ml ; ++++ : 10<sup>6</sup> pfu/ml. pfu : plaque forming units est l'unité de quantification du nombre de phage correspondant au nombre de plaque de lyse formé.

### 2.3.3 Mise au point du crible par phage display

Avant de se lancer dans un crible à grande échelle, il est important de savoir si le phage positif est capable d'être sélectionné dans une population contenant un excès de phage négatif.

Afin de détecter un enrichissement du phage positif par rapport au phage négatif, les inserts contenus dans les phages sont amplifiés par PCR après chaque tour de sélection. L'utilisation d'amorces flanquant les inserts permet une discrimination en fonction de la taille des ADNc. Tout d'abord, nous avons vérifié que l'amplification était possible à partir d'un lysat bactérien. Pour cela le phage positif et le phage négatif sont mélangés à différentes concentrations dans le but d'évaluer les seuils de détection des deux types de phages. Il s'avère que la détection d'un phage en défaut devient difficile à partir d'une différence de 100 fois entre les deux phages (figure 20).



**Figure 20. Détection de l'enrichissement par PCR**

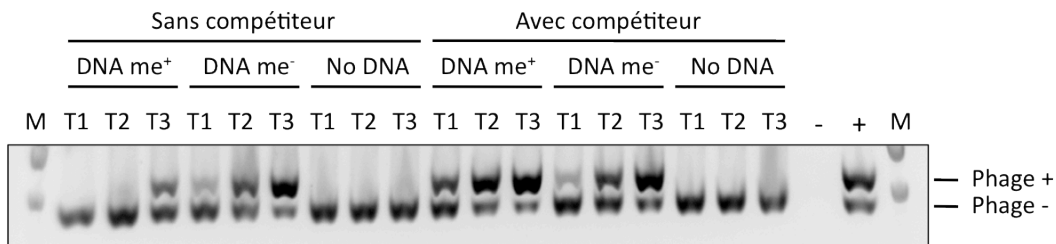
Les phages sont incubés à des concentrations différentes avec des rapports allant de 100/1 à 1/100.

Des expériences d'enrichissement dans lesquelles la quantité de phages négatifs est 1000 fois plus importante que celle du phage positif ont donc été réalisées en présence d'une sonde spécifique correspondant à la séquence du promoteur du gène *Bdnf*, connu pour être lié par MeCP2 (Chen et al. 2003). Nous avons également cherché à déterminer l'effet d'un compétiteur ADN sur la sélection.

Tout d'abord, le phage positif n'est jamais enrichi en absence d'ADN signifiant que les conditions expérimentales et que les plaques streptavidine permettent de réduire au maximum les interactions entre les phages et la matrice. En revanche, nous constatons la sélection du phage positif, au détriment du phage négatif, que ce soit en présence d'ADN méthylé ou d'ADN non méthylé. L'ajout de compétiteur optimise l'enrichissement du phage positif par l'ADN méthylé mais ne permet pas d'éliminer son interaction avec l'ADN non méthylé.

Le crible se révèle donc être efficace quant à la sélection d'un phage positif pour une sonde ADN spécifique mais de façon indépendante de la méthylation et ceci malgré l'utilisation de tampons de liaison ou d'élution différents (Figure 21). Ce résultat peut en partie être expliqué par le fait que MeCP2, bien qu'ayant une affinité importante pour l'ADN méthylé, possède également la capacité de lier l'ADN non méthylé (Klose et al. 2005).

Bien que nous ne soyons pas parvenu à obtenir la sélection du phage positif en fonction de la méthylation de l'ADN, nous estimons, cependant, que les conditions de sélection ne peuvent plus être améliorées à partir de ces expériences. Par conséquent nous avons décidé d'entreprendre un crible génétique, à grande échelle, en utilisant une banque de phages exprimant à leur surface des ADNc de cerveaux humains.



**Figure 21. Le phage positif est enrichi en présence d'ADN par rapport au phage négatif.**

L'enrichissement est détecté par PCR grâce à des amorces flanquant l'ADNc. Le phage positif est incubé avec un excès de  $10^3$  phage négatif en présence d'ADN méthylé, d'ADN non méthylé ou sans ADN. L'effet de l'ajout d'une séquence d'ADN non spécifique en tant que compétiteur est également testé afin d'éliminer les interactions non spécifiques. Avec ou sans compétiteur, le phage positif n'est jamais enrichi après trois tours de sélection. Sans compétiteur, le phage positif est rapidement sélectionné en présence d'ADN non méthylé et seulement après trois tours avec de l'ADN méthylé. L'ajout du compétiteur permet une sélection plus rapide du phage positif sur l'ADN méthylé que sur l'ADN non méthylé.

## 2.4 Résultats des cribles par phage display

Un nombre important de paramètres peut être modifié au cours d'un crible par phage display :

- La séquence cible et sa concentration
- L'utilisation de compétiteur et sa concentration
- Les temps d'incubation
- Les étapes de lavage
- Les tampons de liaison
- Les tampons d'élution
- Les agents bloquants
- Le nombre de tours de sélection

Avant tout démarrage d'un crible par phage display, un paramètre important est à prendre en compte : la représentativité de la banque de phages. La banque d'ADNc de cerveau humain utilisée comprend  $1,5 \times 10^7$  clones différents. Il a été décidé arbitrairement que chaque clone devrait être représenté au moins 100 fois, c'est pourquoi les cribles ont été réalisés à partir de  $2 \times 10^9$  clones au premier tour.

Au cours des différents cribles effectués, deux cibles d'ADN méthylé ont été utilisées. Une première sonde correspond au promoteur du gène *Bdnf*



afin de sélectionner un phage exprimant MeCP2 et ainsi valider définitivement le fonctionnement de la technique. Le désavantage de cette sonde est la spécificité de séquence nous empêchant probablement d'identifier d'autres phages liant l'ADN méthylé. C'est pourquoi nous avons utilisé une sonde artificielle correspondant à un fragment du gène *lacZ*. Cette sonde n'a pas de relevance biologique, mais elle est un contrôle positif *in vitro*, connue pour être liée par Kaiso (Filion et al. 2006). Sa forte densité en CpGs laisse imaginer qu'elle pourrait lier diverses protéines reconnaissant l'ADN méthylé.

L'utilisation de compétiteur semble essentielle pour éliminer les interactions non-spécifiques ne dépendant pas de la méthylation de l'ADN. Une sonde artificielle *LacZ* non méthylée et une séquence de polydIdC ont été utilisées. Une étape de contre-sélection est également envisageable et susceptible de supprimer à la fois les interactions non-spécifiques et les liaisons des phages avec la matrice.

Il est également nécessaire d'étudier l'effet de différents tampons de liaison et des agents bloquants. Pour cette raison, nous avons essayé des solutions de PBS-Tween 1% et de TENN (Table 12). De la BSA 1% ou du lait 5% ont été couplés à ces tampons.

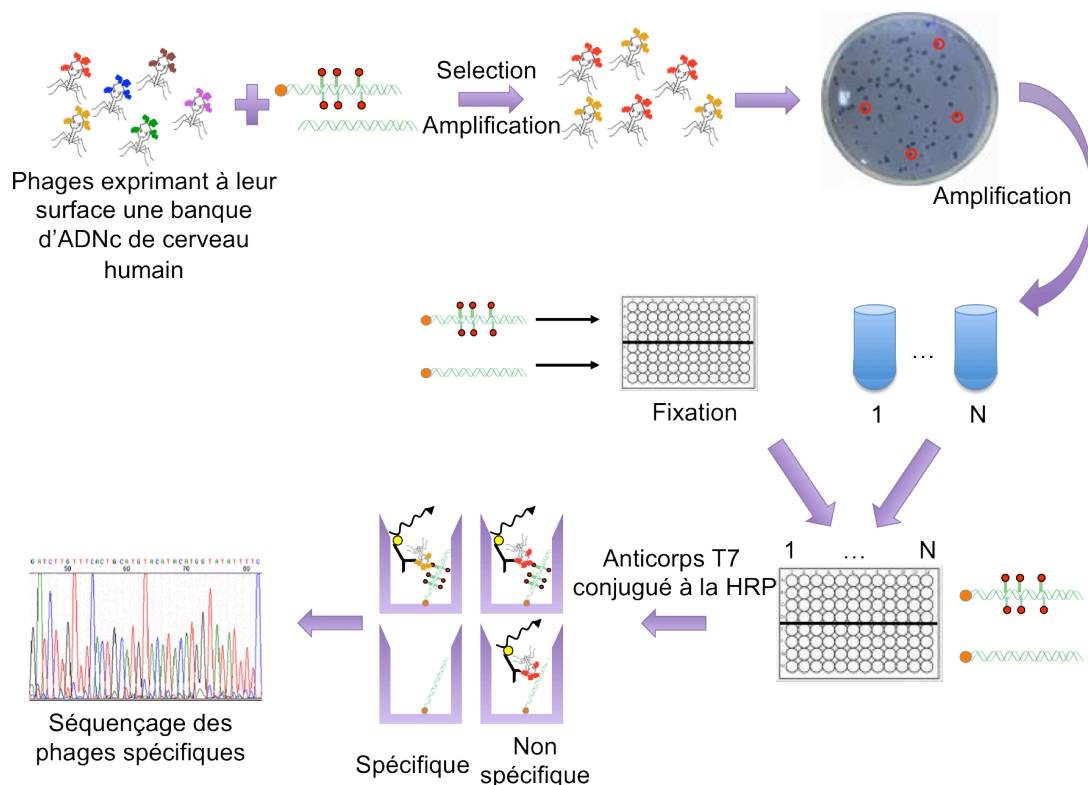
Par ailleurs, il est important de déterminer le nombre de tours de sélection adéquat à effectuer afin d'obtenir une sélection suffisamment importante tout en gardant une certaine hétérogénéité entre les phages et ne pas enrichir en faux-positifs. Pour cela, l'hétérogénéité des phages a été mesurée à chaque tour de crible par PCR sur un échantillon de clones sélectionnés. Quatre à cinq tours de sélection se sont avérés nécessaires pour réaliser un crible.

Enfin, les deux approches permettant la liaison en solution ou après immobilisation de la cible ont été testées. Les conditions de différents cribles réalisés sont listées dans la table 5.

Table 5. Conditions des cribles par phage display.

Crible #	Approche	ADN me <sup>+</sup>	Compétiteur	Contre sélection	Couplage	Sélection	Elution	Tampons	Agents bloquants	Nbre tours
<b>Crible 1</b>	Immobilisation directe	Séquence artificielle 10 pmol	Séquence artificielle 5-10-25-50 pmol	NON	45 min	30 min	SDS 1%	PBST TENN	BSA 1%	4
<b>Crible 2</b>	Immobilisation directe	Séquence <i>Bdnf</i> 10 pmol	polydIdC 1 ug	OUI 15 min	45 min	30 min	SDS 1%	PBST TENN	BSA 1%	5
<b>Crible 3</b>	En solution	Séquence artificielle 50-25-10-5 pmol	Séquence artificielle 5-10-25-50 pmol	NON	20 min	60 min	TEA	PBST TENN	BSA 1%	4
<b>Crible 4</b>	En solution	Séquence artificielle 50 pmol	Séquence artificielle 0-10-20-30-40 pmol	NON	20 min	60 min	TEA	PBST	Lait 5%	5
<b>Crible 5</b>	En solution	Séquence artificielle 50-40-30-20-10 pmol	Séquence artificielle 10-20-30-40-50 pmol	NON	5 min	30 min	TEA	PBST	Lait 5%	5

Lors des cribles 1 à 3, des tests ELISA ont été effectués pour deux raisons. Premièrement, le test ELISA permet de vérifier la véracité des interactions obtenues par phage display et donc de connaître l'enrichissement de faux-positifs. Deuxièmement, ce test reflète directement l'affinité relative d'un phage pour sa cible comparé aux autres clones, nous fournissant ainsi un paramètre supplémentaire afin de choisir les phages à caractériser. La procédure est décrite figure 22 et un exemple de résultat de test ELISA est présenté Table 6.



**Figure 22. Procédure des tests ELISA.**

Les phages sélectionnés par le crible sont amplifiés en infectant des bactéries. Les plaques de lyse ainsi obtenues correspondent à un clone phagique dont l'ADNc code pour une protéine sélectionnée pour interagir avec l'ADN méthylé. Chaque clone est incubé dans une plaque recouverte de Streptavidine en présence d'une sonde méthylée et d'une sonde non méthylée. Un test ELISA est alors réalisé à l'aide d'un anticorps anti-phage T7 conjugué à la HRP. Les phages présentant les plus fortes interactions de manière spécifique avec l'ADN méthylé sont ensuite séquencés.

Les tests ELISA des cribles 1 et 2 montrent que les phages sélectionnés se lient à la cible méthylée avec autant d'affinité qu'à la cible non méthylée (données non présentées). Les résultats de séquençage ne sont alors pas surprenant lorsque l'on constate que les phages enrichis ne sont pas de bons candidats pour se lier à l'ADN méthylé. Dans le cas du crible 1, sur 13 résultats de séquences, seulement 6 clones sont caractérisés comme ayant une localisation nucléaire. Et seul le clone 4, identifié comme étant le gène *Ubtf*, possède un domaine HMG de liaison à l'ADN. Cependant, il est peu probable que la protéine UBTF lie l'ADN méthylé. Premièrement, UBTF est connue pour avoir un effet positif sur la transcription, ce qui va dans le sens contraire à la fonction primaire de la méthylation de l'ADN qui a généralement un rôle négatif concernant l'expression des gènes. Deuxièmement, les résultats de séquençage des autres clones indiquent que les conditions du crible ont permis essentiellement la sélection de faux-positifs, ce qui est confirmé par le test ELISA. Il est donc très peu probable que l'enrichissement de UBTF soit spécifique de la méthylation de l'ADN.

Lors du crible 3, les phages de la banque ont été sélectionnés selon les conditions décrites en table 5. De façon indépendante, les tampons PBST et TENN ont été utilisés et les phages enrichis après 4 tours ont été testés pour leur capacité à lier spécifiquement l'ADN méthylé par ELISA. La comparaison entre les deux types de tampons permet de dire que le PBST semble plus adapté pour l'enrichissement de phages positifs que le TENN. En effet, les résultats obtenus avec le PBST montrent que 11 phages sélectionnés sur 23 ont une affinité plus importante pour l'ADN méthylé que pour le non méthylé, contre 2 sur 23 dans le cas du TENN. Les 13 phages surlignés en jaune et orange ont alors été envoyés à séquencer. Les résultats de séquençage sont présentés Table 9.

Le test ELISA du crible 3 présente des différences nettes d'affinité selon l'état de méthylation de la sonde. Néanmoins, comme pour les cribles 1 et 2, le crible 3 ne présente pas de candidats intéressants à l'exception de ARID5B connu pour lier la séquence CCGCCC (Tu et al. 2008).

		1	2	3	4	5	6	7	8	9	10	11	12	
ADN méthylé	A	0,23	0,19	0,18	0,22	0,24	0,25	0,16	0,24	0,21	0,27	0,18	0,32	PBST
	B	0,34	0,18	0,16	0,29	0,21	0,17	0,18	0,17	0,16	0,23	0,19	0,17	
	C	0,18	0,17	0,17	0,16	0,17	0,18	0,17	0,16	0,17	0,19	0,17	0,22	TENN
	D	0,19	0,18	0,18	0,17	0,17	0,22	0,18	0,19	0,18	0,17	0,19	0,17	
ADN non méthylé	E	0,17	0,17	0,17	0,18	0,19	0,20	0,18	0,27	0,18	0,18	0,17	0,23	PBST
	F	0,25	0,16	0,16	0,31	0,18	0,17	0,18	0,18	0,17	0,21	0,17	0,16	
	G	0,18	0,16	0,16	0,16	0,17	0,16	0,17	0,16	0,17	0,17	0,17	0,19	TENN
	H	0,17	0,18	0,16	0,16	0,14	0,16	0,16	0,16	0,16	0,16	0,17	0,22	

**Table 6. ELISA du crible 3.**

La plaque est séparée en deux avec une partie des puits liée par un ADN méthylé et une autre partie par un ADN non méthylé. Chaque clone est incubé avec à la fois avec l'ADN méthylé et l'ADN non méthylé. Deux tampons (PBST et TENN) ont été utilisés séparément lors du crible dans des conditions identiques, l'effet de ces deux tampons est ici analysé. En rouge, les puits de références ne contenant pas de phages et permettant de déterminer le bruit fond du test ELISA. Surlignés en jaune, les puits contenant les phages se liant plus fortement à l'ADN méthylé que non méthylé. Surlignés en orange, les phages ayant l'affinité la plus importante. Surlignés en vert, un phage présentant une plus grande affinité pour l'ADN non méthylé que méthylé.

**Table 7. Résultats du crible 1.**

Crible 1	Séquence	Région interaction	Localisation	Fonction
1	Ring Finger Protein 11 (RNF11)	3'UTR	Nucléaire Cytoplasmique	Complexe ubiquitin-ligase
2	Oxysterol-Binding Protein-Like 1A (OSBPL1A)	Exon	Membranaire	Récepteur phospholipide
3	Fas Apoptotic Inhibitory Molecule (FAIM)	Exon	Cytoplasmique	Apoptose
4	Upstream Binding Transcription Factor (UBTF)	Exon	Nucléaire	Liaison ADN (domaine HMG) Expression ADNr
5	Zinc finger CCCH-type Containing 3 (ZC3H3)	Intron	Nucléaire	/
6	Potassium Channel Tetramerisation Domain 4 (KCTD4)	Promoteur	cytoplasmique	/
7	Transcription Elongation Factor A (SII)-like 6 (TCEAL6)	Exon	Nucléaire	Transcription? Aucun domaine particulier
8	Tyrosine 3-Monooxygenase/Tryptophan 5-Monooxygenase (YWHAH)	ARN non-codant	Cytoplasmique	Transduction de signal
9	Zinc finger CCCH-type Containing 3 (ZC3H3)	Intron	Nucléaire	/
10	UNC13C	Intron	Cytoplasmique	Rôle dans l'exocytose
11	Ornithine Decarboxylase antizyme 2 (OAZ2)	Exon	Nucléaire	Métabolisme
12	Ubiquitin Specific Peptidase 54 (USP54)	Exon	Cytoplasmique	Inactive
13	Dynein, Axonemal, Heavy Polypeptide 6	Intron	Cytoplasmique	Activité moteur des microtubules

Nous pouvons néanmoins expliquer les résultats de séquençage par les faibles différences d'affinités relatives obtenues entre les sondes indiquant probablement qu'elles sont non significatives. Par ailleurs, le signal généré est certainement négatif au vu des valeurs relativement faibles des intensités des signaux. A posteriori, il semble évident que l'utilisation du phage positif, MeCP2, en tant que contrôle positif du test ELISA était nécessaire afin d'évaluer l'affinité des phages sélectionnés.

Par la suite, la BSA, utilisée en tant qu'agent bloquant, a été remplacée par du lait lors du crible 5, et les temps d'incubation ont été réduits afin de diminuer les interactions faibles et non-spécifiques. Les résultats de séquençages sont présentés Table 10. Contrairement aux cribles précédents, 8 protéines sur 12 sont nucléaires dont 6 capables de lier un acide nucléique. Parmi elles, CHD3, HMGB1, NEUROD2 et KLF6 se présentent comme des candidats potentiels.

La protéine CHD3, également appelée Mi2, est caractérisée pour faire partie du complexe de remodelage de la chromatine NuRD qui comprend aussi HDAC1 et MBD3 (Tong et al. 1998, Zhang et al. 1999). Le complexe NuRD s'associe aux centrosomes, régions fortement méthylées, pour réguler l'intégrité de cette structure (Sillibourne et al. 2007). Il a été proposé que le complexe NuRD soit recruté via la reconnaissance de l'ADN méthylé. Cependant, MBD3 ne possède pas la capacité de lier la méthylcytosine. Il serait donc intéressant de savoir si CHD3 se lie à l'ADN méthylé et si la délétion de CHD3 inhibe le recrutement du complexe NuRD à la chromatine. Toutefois, Zhang et ses collègues ont montré par des expériences de retard sur gel et par des GST pull-down que MBD2 interagit avec et recrute le complexe NuRD à l'ADN (Zhang et al. 1999). Il reste néanmoins envisageable que la protéine CHD3 puisse lier l'ADN méthylé.

**Table 8. Résultats du crible 2.**

Crible 2	Séquence	Région interaction	Localisation	Fonction
1	Nucleolar and coiled-body phosphoprotein 1 (NOLC1)	Exon	Nucléolaire	Nucléogénèse
2	Nucleolar and coiled-body phosphoprotein 1 (NOLC1)	Exon	Nucléolaire	Nucléogénèse
3	Séquence non codante	Intergénique	/	/
4	CDC42-binding protein kinase alpha (CDC42BPA)	Intron	Cytoplasmique	Kinase
5	RNA Binding Motif Protein 39 (RBM39)	Exon	Nucléaire	Liaison ARN
6	Fatty Acyl CoA Reductase 1 (FAR1)	3'UTR	Membranaire	Métabolisme acyl-CoA
7	Nucleolar and coiled-body phosphoprotein 1 (NOLC1)	Exon	Nucléolaire	Nucléogénèse
8	DDHD domain containing 2 (DDHD2)	Exon	Cytoplasmique	Phospholipase
9	Nucleolar and coiled-body phosphoprotein 1 (NOLC1)	Exon	Nucléolaire	Nucléogénèse
10	Nucleolar and coiled-body phosphoprotein 1 (NOLC1)	Exon	Nucléolaire	Nucléogénèse
11	Nucleolar and coiled-body phosphoprotein 1 (NOLC1)	Exon	Nucléolaire	Nucléogénèse
12	Nucleolar and coiled-body phosphoprotein 1 (NOLC1)	Exon	Nucléolaire	Nucléogénèse
13	ND	/	/	/
14	Nucleolar and coiled-body phosphoprotein 1 (NOLC1)	Exon	Nucléolaire	Nucléogénèse
15	Nucleolar and coiled-body phosphoprotein 1 (NOLC1)	Exon	Nucléolaire	Nucléogénèse

**Table 9. Résultats du crible 3.**

Crible 3	Séquence	Région interaction	Localisation	Fonction
1	Scaffold Attachment Factor B (SAFB)	Intron	Nucléaire	Liaison à l'ADN attaché à la matrice
2	Nucleolar and Spindle Associated Protein 1 (NuSAP1)	Intron	Cytoplasmique Nucléaire	Organisation microtubule
3	Séquence non codante	Intergénique	/	/
4	Family with sequence similarity 43, member B (FAM43B)	Exon	ND	ND
5	Chromosome 20 Open Reading Frame 30 (C20orf30)	Exon	ND	ND
6	RNA binding motif protein 8B (RBM8B)	Exon	Nucléaire	Liaison ARN
7	baculoviral IAP repeat-containing 6 (BIRC6)	Exon	Cytoplasmique	Ubiquitin-ligase Rôle anti-apoptotique
8	AT rich interactive domain 5B (ARID5B)	3'UTR	Noyau	Liaison à l'ADN Régulateur transcriptionnel
9	succinate dehydrogenase complex, subunit A (SDHA)	Exon	Mitochondrie (Membrane)	Chaîne respiratoire
10	MEG3	Non-codante	Nucléaire	ARN non-codant Gène soumis à l'empreinte
11	Dual specificity tyrosine-phosphorylation-regulated kinase 1A (DYRK1A)	3'UTR	Nucléaire	Kinase
12	Non déterminée	/	/	/
13	HtrA serine peptidase 1	Intron	Sécrétée	Protéase

La protéine NEUROD2, en régulant les promoteurs spécifiques des neurones, est importante pour leur différenciation. Elle possède un motif basic helix-loop-helix (bHLH) caractéristique de certains facteurs de transcription qui reconnaît une séquence canonique CACGTG nommée E-box (Massari et Murre 2000). De manière similaire à d'autres facteurs bHLH, NEUROD2 lie directement les E-box de certains gènes cibles et stimule leur expression (Lin et al. 2005). Récemment, des expériences de retard sur gel ont montré que la méthylation de la cytosine réduisait la liaison de protéines bHLH (Takahashi et al. 1998 ; Kinoshita et al. 2004). Il est donc probable que NEUROD2 se comporte comme les autres protéines de la famille bHLH. Cependant, les protéines à doigts de zinc sont un contre-exemple parfait d'une famille dans laquelle certains membres peuvent lier l'ADN méthylé et d'autres non.

KLF6 est un facteur de transcription de la famille des protéines à doigts de zinc de type Krüppel-like qui joue un rôle dans divers processus biologique dont le développement, la différenciation, la prolifération et l'apoptose (Bieker et al. 2001). Il a été récemment montré, par des expériences de retard sur gel et de ChIP, que la méthylation de l'ADN diminue la liaison de KLF6 sur le promoteur du gène TFPI-2 (Guo et al. 2007). KLF6 a donc été enrichie de façon non spécifique au cours du crible.

La protéine HMGB1 a été identifiée comme liant et modifiant la conformation de l'ADN via les domaines HMG. HMGB1 facilite le recrutement de complexes nucléoprotéiques et régulent ainsi l'expression des gènes (Thomas et Travers 2001). HMGB1 semble donc être un candidat intéressant pour lier l'ADN méthylé. En 2004, D. Wang a montré au cours de sa thèse par des expériences de retard sur gel que la méthylation de l'ADN réduit l'interaction du domaine HMG avec l'ADN (non publié). Bien que l'affinité pour l'ADN méthylé semble plus faible, l'interaction entre HMGB1 et la sonde méthylée persiste. La question sur la capacité de HMGB1 à lier l'ADN méthylé n'est cependant pas résolue et des expériences *in vivo* de colocalisation avec les chromocentres dans des cellules de souris ainsi que des expériences *in vitro* de retard sur gel sont nécessaires afin d'adresser cette énigme.



**Table 10. Résultats du crible 4.**

Crible 4	Séquence	Région interaction	Localisation	Fonction
1	Neurogranin (NRGN)	Gène entier	cytoplasmique	Transduction de signal
2	Séquence non codante	Intergénique	/	/
3	Neurogranin (NRGN)	Gène entier	cytoplasmique	Transduction de signal
4	Cell Adhesion Molecule 1 (CADM1)	Intron	Cystoplasmique Membranaire	Adhésion cellulaire
5	Séquence non codante (identique au clone 2)	Intergénique	/	/
6	Neurogranin (NRGN)	Gène entier	cytoplasmique	Transduction de signal
7	Neurogranin (NRGN)	Gène entier	cytoplasmique	Transduction de signal
8	Formin binding Protein 1 (FBNP1)	Exon	Cytoplasmique Membranaire	Endocytose
9	Neurogranin (NRGN)	Gène entier	cytoplasmique	Transduction de signal
10	Pre-B-cell Leukemia Homeobox 1 (PBX1)	Intron	Nucléaire Cytoplasmique	Liaison ADN via Homeobox
11	Synaptosomal-Associated Protein 25 (SNAP25)	Exon	Cytoplasmique Membranaire	Régulation de la sécrétion des neurotransmetteurs
12	Suppressor of Cancer Cell Invasion 1 (SCA1)	Intron	Cytoplasmique Nucléaire	Transduction de signal
13	Signal Recognition Particle 9kDa (SRP9)	3'UTR	Cytoplasmique	Ribonucléoprotéine

**Table 11. Résultats du crible 5.**

Crible 4	Séquence	Région interaction	Localisation	Fonction
1	Cold Inducible RNA Binding Protein (CRBP1)	Exon	Nucléaire	Liaison ARN Réponse aux stress
2	Chromodomain Helicase DNA Binding Protein 3 (CHD3)	Exon	Nucléaire	Liaison ADN Doigts de zinc PHD Chromodomaine Composant du complexe NuRD
3	Mitochondrial Ribosomal Protein L13 (MRPL13)	Intron	Mitochondrie	Ribonucléoprotéine
4	Beta-Adrenergic Receptor Kinase 1 (ARBK1)	Exon	Cytosolique	Kinase
5	High-Mobility Group Box 1 (HMGB1)	Exon	Nucléaire	Liaison ADN Domaine HMG Recombinaison V(D)J
6	Neurogenic Differentiation 2 (NeuroD2)	Exon	Nucléaire	Liaison ADN Différenciation Régulation transcription
7	Coiled-Coil Domain Containing 104 (CCDC104)	Exon	Nucléaire	/
8	Nucleolar and coiled-body phosphoprotein 1 (NOLC1)	Exon	Nucléolaire	Nucléogénèse
9	Séquence non codante	Intergénique	/	/
10	Séquence non codante	Intergénique	/	/
11	Kruppel-Like Factor 6 (KLF6)	Exon	Nucléaire	Liaison ADN Doigts de zinc
12	Splicing Factor 3a (SF3A3)	Exon	Nucléaire	Epissage Doigts de zinc

## 2.5 Discussion

Nous avons cherché à développer un crible par phage display afin d'identifier des protéines liant l'ADN méthylé. Nous avons procédé par étapes afin d'optimiser au mieux le crible en recherchant, tout d'abord, la matrice adéquate pour limiter les interactions non-spécifiques. Les plaques recouvertes de streptavidine se sont avérées être un meilleur support que les billes magnétiques. Ensuite, dans des conditions simples, nous avons mis au point les conditions permettant l'enrichissement d'un phage positif sur un phage négatif présent en excès. Nous avons alors réalisé différents cribles en variant divers facteurs comme les tampons de liaison, les temps d'incubation, l'ajout de compétiteur, la stringence des lavages et les séquences cibles. Les cribles 1 à 4 ont montré leur faible sensibilité conduisant à l'enrichissement de faux-positifs. Un nombre important de protéines cytoplasmiques et transmembraires ont été sélectionnées reflétant de mauvaises conditions expérimentales. Enfin, lors du crible 5, les résultats de séquençage montrent une sélection enrichie en facteur de transcription. Cela démontre que nous sommes capables d'enrichir des protéines liant les acides nucléiques et donc potentiellement un ADN méthylé. Néanmoins, comme nous avons pu le voir précédemment, les protéines, à l'exception peut-être de CHD3 et de HMGB1, ne représentent pas de bons candidats à caractériser ultérieurement. Par ailleurs, un nombre restreint de phages a été séquençé, ne nous permettant pas de déterminer la diversité des phages enrichis ainsi que la représentativité de chaque clone au sein de la population sélectionnée. Nous avons donc la nécessité de séquencer plus de clones.

Malheureusement, notre incapacité à identifier une MBP caractérisée ne nous permet pas de valider cette approche comme étant un moyen de trouver de nouvelles MBPs, malgré l'utilisation d'une banque de cerveau présentant une expression abondante de MeCP2 dans les neurones. Lors du crible 5, nous avons utilisé une sonde méthylé riche en G-C. Or, il a été démontré que la liaison de MeCP2 à ses cibles nécessitait des séquences adjacentes riches en A-T (Klose et al. 2005).

Un paramètre important à prendre également en considération est le caractère *in vitro* de ces expériences. Il est donc concevable que certaines MBPs préfèrent lier l'ADN méthylé dans un contexte chromatinien. Récemment, un nouvel outil a été développé par les laboratoires du Dr Kouzarides et du Dr Mann afin d'identifier des facteurs capables de reconnaître la méthylation de l'ADN et/ou des histones dans un contexte de nucléosomes (Bartke et al. 2010), à partir d'une méthode d'analyse protéomique appelée SILAC (Ong et al. 2002). Ils ont été en mesure de montrer un enrichissement de MBD2, MBD4, MeCP2 et UHRF1 à l'ADN méthylé et de nombreuses protéines non caractérisées pour lier l'ADN méthylé ont été obtenues (Bartke et al. 2010). Ces protéines sont donc de bons candidats pour des analyses approfondies.

Finalement, d'autres approches peuvent être envisagées pour identifier de nouvelles MBPs. Récemment, un crible par simple hybride a été mis au point (Feng et al. 2004) et les auteurs ont pu identifier MBD2 et MBD4 au cours de leurs cribles validant ainsi la fonctionnalité de leur approche (Feng et al. 2010). Néanmoins, les autres protéines identifiées s'avèrent des faux-positifs et ils n'ont pas été en mesure de découvrir de nouvelles MBPs.

## 2.6 Matériels et méthodes

### Hybridation des sondes

Les oligonucléotides ont été synthétisés, méthylés et biotinylés par *Sigma-Aldrich*. 10 µl de chaque oligonucléotides à 10 µM est incubé avec 10 µl de solution de NaCl 0,5 M et 70 µl d'eau afin d'obtenir une concentration finale de 1µM. La réaction est alors incubée dans une machine PCR à 95° C pendant 5 min puis la température est graduellement diminuée jusqu'à 25° C pendant 45 min.

### Titration

La titration permet de déterminer le nombre de phage présent dans les échantillons après sélection ou amplification. La technique repose sur l'incubation d'échantillons de phages dilués avec la souche de bactéries hôtes 5615 dans de l'agarose en solution destinée à être étalé sur des boîtes

de Pétri LB agar. Les phages provoquent alors une plage de lyse sur la couche de bactérie en croissance. L'agarose est chauffé à 90° C afin de le liquifier puis est maintenu dans cet état à 50° C. Les phages sont dilués à diverses concentrations dans du LB et 250 µl d'une culture de bactéries 5615 est incubé avec 100 µl de la solution de phages dilués. En parallèle, 100 µl d'IPTG 100 mM sont ajoutés à la solution d'agarose pour ensuite mélanger 2,5 ml d'agarose avec la solution de phages-bactéries. Le mélange est alors étalé sur des boîtes de Pétri LB agar à 37° C pendant 3-4 heures. Les plages de lyses sont finalement comptées et la concentration de phages est déterminée par la quantité de plages de lyses formées par unité de volume.

### Clonage de MeCP2 dans un phage

La région codante pour le domaine MBD de MeCP2 a été amplifiée par PCR. La polymérase Phusion (*Finnzymes*) est utilisée pour réaliser les amplifications par PCR des clones. Le produit d'amplification est ensuite purifié par passage sur mini colonne *NucleoSpin Extract II* (*Macherey-Nagel*) et repris dans 50 µl de tampon d'élution AE (5mM Tris/HCl, pH 8.5). 1 µg de produit purifié est digéré par les enzymes EcoR1 et Xho1 pendant 2 heures à 37°C puis purifié sur mini colonne *NucleoSpin Extract II* et repris dans 30 µl de tampon d'élution AE. La fusion à la capside 10B est obtenue en clonant le produit digéré dans le vecteur T7Select10-3b (*Novagen*). 2 µg de T7Select10-3b sont digérés pendant 1 heure et demi à 37°C puis purifiés par mini colonne *NucleoSpin Extract II* et repris dans 30 µl de tampon d'élution AE. 150 ng de vecteur sont combinés avec 3 fois d'excès molaire d'insert puis le volume est ajusté à 15 µl avec du TE (10 mM Tris, pH 8 et 1 mM EDTA). 1,5 µl de tampon 10X Ligation Buffer (*Novagen*) sont ajoutés ainsi que 1,5 µl de T4 DNA Ligase (*Novagen*) et 1,5 de DTT 100 mM. La réaction de ligation est incubée sur la nuit à température ambiante. 15 µl de produit de ligation sont ensuite incubés avec 10 µl de tampon T7Select Packaging Extract (*Novagen*) permettant une réaction d'encapsidation *in vitro*. Les phages sont alors amplifiés et l'insertion de MeCP2 est vérifiée par PCR puis séquençage.

### Blocage de la matrice

Cent  $\mu\text{l}$  de billes magnétiques recouvertes de streptavidine (Dynabeads M-280 Streptavidin, *Invitrogen*) sont lavées dans 1 ml de PBS puis resuspendues dans 1 ml de PBS. Les billes sont alors bloquées avec de la BSA à 0,5 mg/ml pendant une heure sur roue.

Lors du crible 5, des tubes de 1,5 ml et la plaque streptavidine (Reacti-bind Neutravidin coated plate, *ThermoScientific*) sont incubés avec du PBS Tween 1% et 5% lait pendant une heure.

#### Sélection du crible 4

$2 \times 10^8$  phages sont incubés dans 200  $\mu\text{l}$  de PBST 1%, 5% lait dans des tubes de 1,5 ml, préalablement bloqués, en présence de 50 pmol de sonde méthylée et de 10 pmol de compétiteur pendant 30 min. La solution est alors récupérée et incubée dans les plaques streptavidine afin d'immobiliser les complexes phages-ADN pendant 5 min. Une série de 15 lavages au PBST est réalisée puis les phages liés sont élués à l'aide de 100  $\mu\text{l}$  de TEA 100 mM pendant une minute. La réaction d'éluion est alors arrêtée par l'ajout de 100  $\mu\text{l}$  de HCl 250 mM. L'éluat de phage est ensuite amplifié afin d'effectuer d'autres tours de sélection. Cinq tours de sélection ont ainsi été réalisés. La quantité de sonde méthylée a graduellement été réduite au cours des tours tandis que celle du compétiteur a été augmentée (voir Table 5).

#### Test ELISA

La plaque streptavidine (Reacti-bind Neutravidin coated plate, *ThermoScientific*) est bloquée avec du PBST 1%, 5% lait pendant une heure à température ambiante. Cinquante pmol d'oligonucléotides biotinylés en solution dans du PBST 1%, 5% lait sont incubés dans les puits pendant 30 min. Ensuite, 100  $\mu\text{l}$  de lysat de phage sont ajoutés dans les différents puits pendant 30 min. Les phages non liés sont alors éliminés par une succession de 5 lavages avec du PBST 1%, 5% lait puis un lavage avec du PBS. Une solution d'anticorps T7 couplé à la HRP (*Novagen*) est préparée au dix-millième dans 100  $\mu\text{l}$  de PBS puis ajouté aux puits pendant une heure. Afin d'éliminer les anticorps non liés, 5 lavages au PBST 1%, 5% lait et un lavage au PBS ont été effectués. Cent  $\mu\text{l}$  de solution de TMB (*Sigma-Aldrich*) sont ajoutés pendant 20 min pour détecter la HRP de l'anticorps T7 (*Novagen*). La

réaction est alors arrêtée par l'ajout de 100 µl de HCl 250 mM. L'absorbance est alors obtenue à l'aide d'un luminomètre GloMax MultiDetection System (Promega).

### Tampons

Les tampons utilisés sont listés table 12

**Table 12. Tampons**

BW 1X	TENN 1X	TENN 1X	PBST
5mM Tris-HCl pH7.5 1M NaCl 0.5 mM EDTA	20mM Tris-HCl pH7.5 1mM EDTA 100mM NaCl 0.1% NP-40	20mM Tris-HCl pH7.5 1mM EDTA 200mM NaCl 0.1% NP-40	PBS 1X Tween 1% Lait 5%

### Sondes

Les sondes utilisées au cours des différents cribles par phage display sont listées table 13.

**Table 13. Sondes.**

Oligo	Nombre de CG	Méthylation	Biotinylation	Séquence	Commentaire
MJ45	1	+	+	CAAAGCATGCAATGCCCTGGAAC <b>CG</b> GAATTCTTCT AATAAAAGATGTATCATTTTAAATGC	Bdnf
MJ46	1	+	-	GCATTTAAATGATACATCTTTTATTAGAAGAATTC <b>CG</b> TTCCAGGGCATTGCATGCTTTG	Bdnf
MJ47	1	-	-	CAAAGCATGCAATGCCCTGGAAC <b>CG</b> GAATTCTTCTAATAA AAGATGTATCATTTTAAATGC	Bdnf
MJ48	1	-	-	GCATTTAAATGATACATCTTTTATTAGAAGAATTC <b>CG</b> TT CCAGGGCATTGCATGCTTTG	Bdnf
MJ58	12	-	-	ATT <b>CGCGCG</b> CTGGCTAC <b>CGGCG</b> ATGAG <b>CGAACGCG</b> TAA <b>CGCG</b> AATGGTGCAG <b>CGCG</b> ATCG	LacZ
MJ59	12	-	-	<b>CGATCGCG</b> CTGCACCATTC <b>CGCG</b> TTAC <b>CGCG</b> TT <b>CG</b> TCATC GCCGGTAGCCAGCG <b>CGCG</b> GAAT	LacZ
MJ60	12	+	+	ATT <b>CGCGCG</b> CTGGCTAC <b>CGGCG</b> ATGAG <b>CGAACGCG</b> TAA <b>CGCG</b> AATGGTGCAG <b>CGCG</b> ATCG	LacZ
MJ61	12	+	-	<b>CGATCGCG</b> CTGCACCATTC <b>CGCG</b> TTAC <b>CGCG</b> TT <b>CG</b> TCATC GCCGGTAGCCAGCG <b>CGCG</b> GAA	LacZ



## Chapitre 3 : approche basée sur la littérature

### 3.1 Introduction

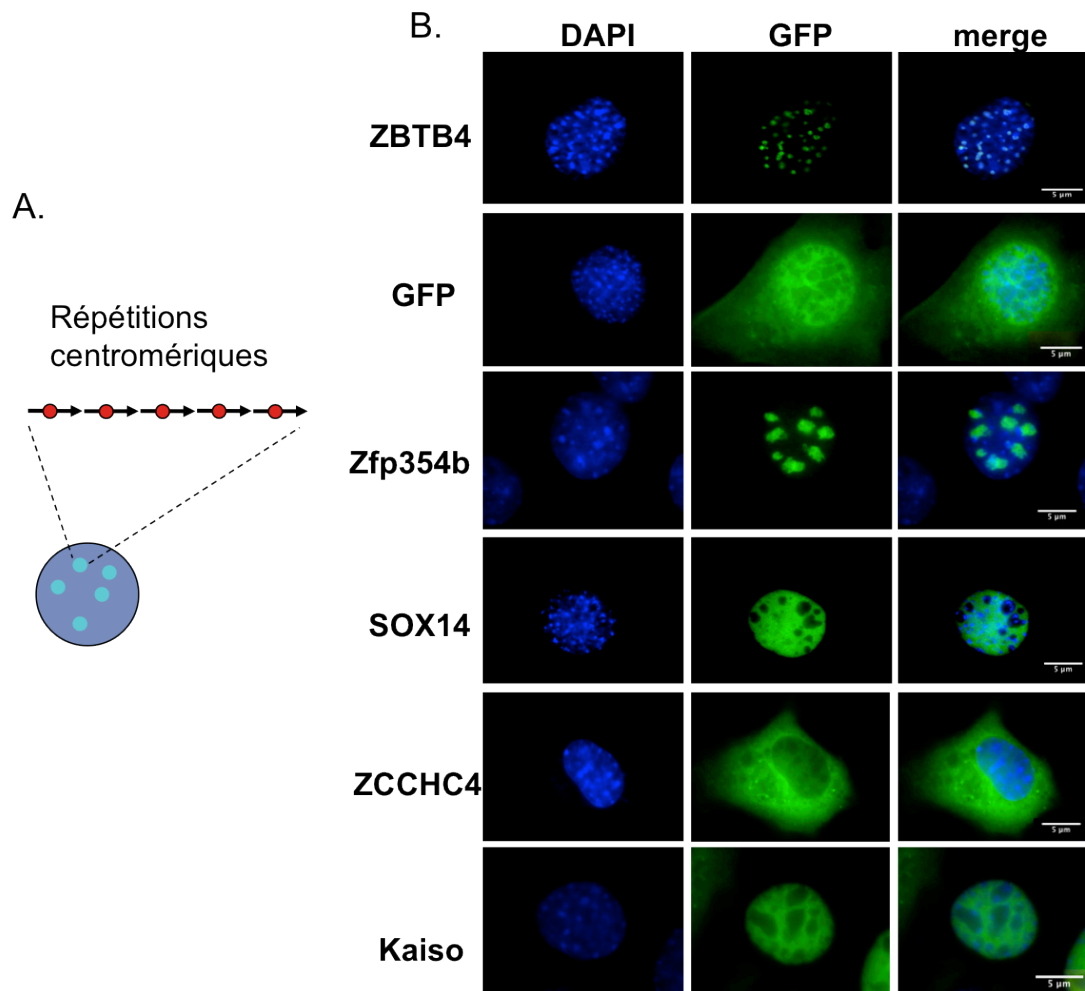
Lors de la transformation tumorale, de nombreux promoteurs deviennent anormalement méthylés (voir Partie 1, chapitre 3.8.3). Les mécanismes reliant transformation et méthylation sont mal connus et, pour les éclaircir, un article récent a utilisé une approche de crible fonctionnel d'ARN interférence (Gazin et al. 2007). Les auteurs ont recherché les protéines dont la déplétion empêche la méthylation du promoteur Fas, qui est normalement observée lorsque des cellules sont transformées par l'oncogène K-ras.

Une trentaine de protéines a été obtenue. On y trouve, comme on pourrait l'attendre, l'ADN méthyltransférase DNMT1. Parmi les autres on peut espérer trouver des protéines liant l'ADN méthylé et favorisant l'hétérochromatinisation du promoteur Fas. Leur absence pourrait être la cause de la ré-expression du promoteur, et par conséquent de sa déméthylation.

Beaucoup des protéines candidates sont non caractérisées, mais certaines présentent un intérêt particulier du fait qu'elles contiennent des domaines connus pour permettre la liaison à l'ADN. Plus encore, certaines des protéines candidates ont été étudiées par immunoprécipitation de chromatine, et sont présentes sur le promoteur Fas méthylé, mais non sur le promoteur Fas déméthylé. Ce sont donc d'excellents candidats pour notre recherche.

Nous avons donc étudié les protéines identifiées par Gazin et al. pour leur rôle dans le maintien de la méthylation du promoteur Fas, en nous restreignant aux protéines dont la séquence suggérait qu'elles pouvaient lier l'ADN. La liste de ces protéines, ainsi que leurs caractéristiques connues, est présentée dans la table.





**Figure 23. Localisation des protéines candidates en cellules murines.**

**A :** Organisation du noyau des cellules de souris. Les noyaux présentent des régions denses au DAPI appelées chromocentres. Elles correspondent aux répétitions péri-centromériques où l'ADN est fortement méthylé. **B :** ZBTB4, la GFP, Kaiso et les protéines candidates ont été exprimés en fusion à la GFP dans les cellules MEF  $p53^{-/-}$ . Le contrôle positif : ZBTB4 colocalise au niveau des chromocentres tandis que la GFP seule est diffuse dans toute la cellule. Malgré la capacité de Kaiso à lier l'ADN méthylé, cette dernière n'est pas localisée aux chromocentres. Zfp354b semble colocaliser avec les nucléoles alors que zcchc4 est cytoplasmique. Sox14 est diffuse dans le noyau.

## 3.2 Résultats

### 3.2.1 Les réactifs générés

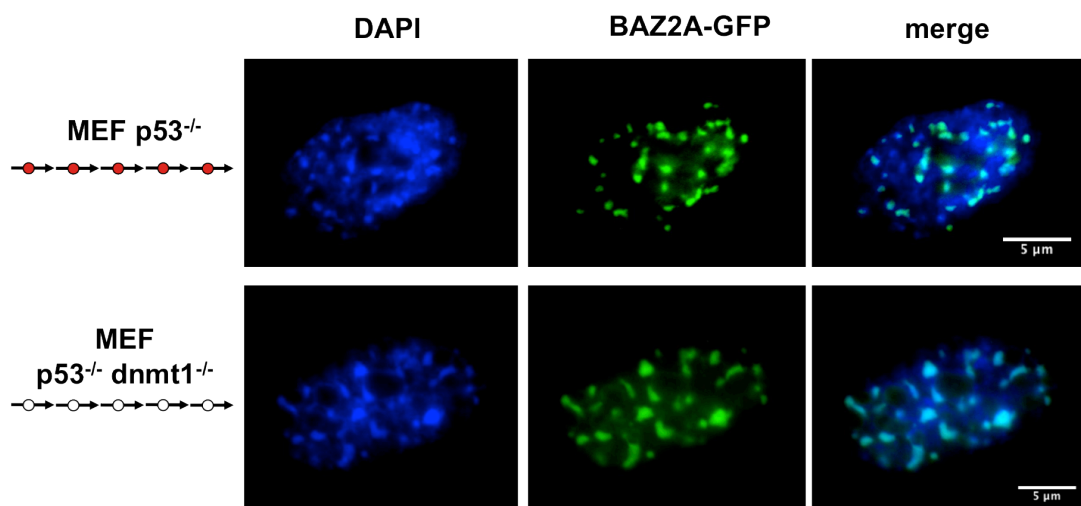
Notre stratégie a été d'utiliser en parallèle des approches *in vivo* et *in vitro*. Pour cela, les protéines candidates ont été clonées :

- Soit dans un plasmide pEGFP-C2 permettant l'expression des protéines en fusion à la GFP, pour localiser les candidats *in vivo*. La localisation subcellulaire est en effet un indice sur leur possible spécificité pour l'ADN méthylé.
- Soit dans un plasmide pCITE-4a afin de synthétiser les protéines *in vitro* en lysat de réticulocyte de lapin. Ces protéines servent ensuite dans des expériences de retard sur gel.

### 3.2.2 Résultats des expériences de colocalisation

Les cellules murines présentent des foyers nucléaires qui se colorent fortement au DAPI et qu'on appelle chromocentres. De nombreuses protéines liant l'ADN méthylé colocalisent avec ces structures. Notons qu'il y a des exceptions, notamment la protéine Kaiso qui ne s'accumule pas aux chromocentres, bien qu'elle lie l'ADN méthylé *in vitro*. Nous reviendrons sur cet état de fait dans la discussion. Notons que les cellules humaines ont une organisation différente, et ne présentent pas de chromocentres. C'est pour cette raison expérimentale que nous avons utilisé des cellules de souris.

Il existe des cellules murines dans lesquelles les répétitions sont déméthylées. Ces cellules sont très utiles pour déterminer si une localisation aux chromocentres dépend de la présence d'ADN méthylé, ou si au contraire elle est maintenue même lorsque l'ADN est déméthylé. Ces cellules ont été générées par recombinaisons homologues des deux allèles de DNMT1. Ces cellules ont une propension forte à rentrer en apoptose (Jorgensen et al. 2004). Elles sont stabilisées par une mutation homozygote de p53. Les deux types de cellules transfectées sont donc les cellules de référence p53<sup>-/-</sup> et les cellules déméthylées dnmt1<sup>-/-</sup>/p53<sup>-/-</sup>. Nous avons vérifié que la mutation de p53 ne modifiait pas les résultats obtenus en répétant les expériences dans des cellules 3T3, qui expriment p53 (données non montrées ici).



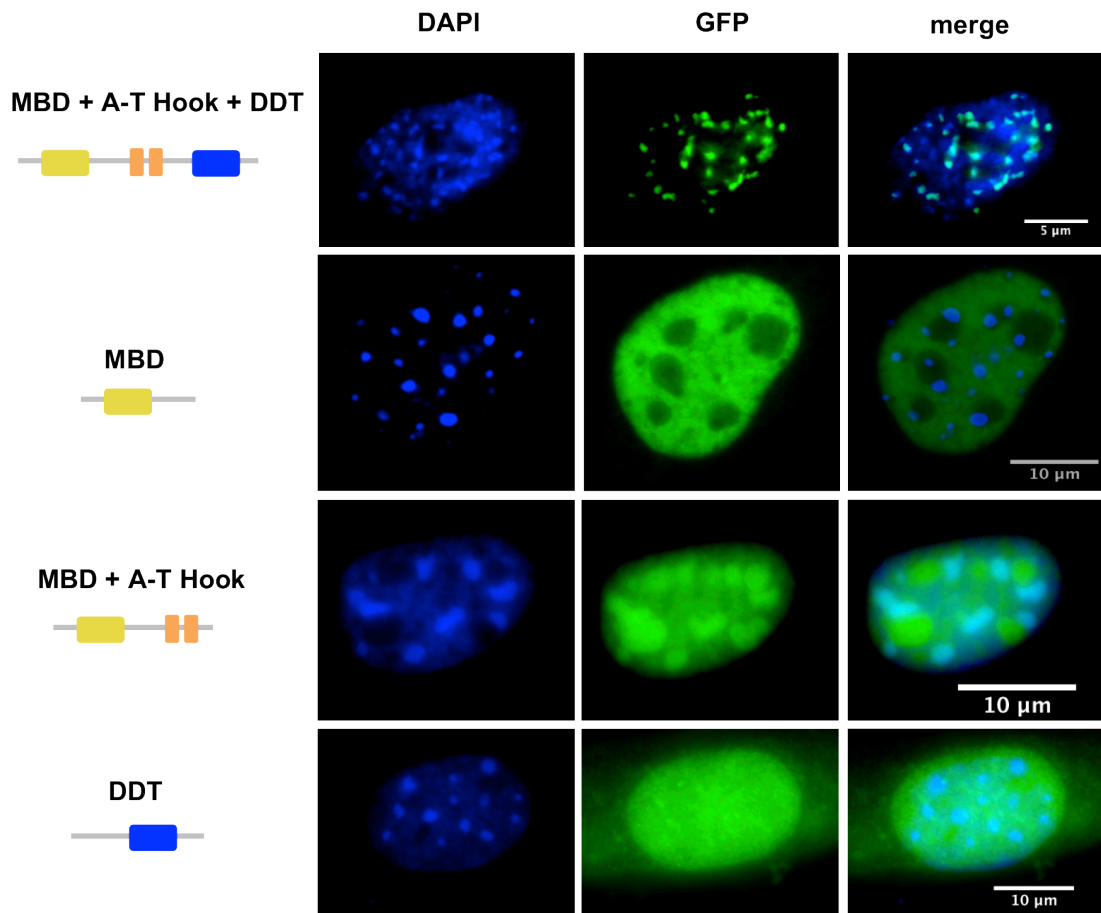
**Figure 24. BAZ2A colocalise avec les chromocentres indépendamment de la méthylation.**

BAZ2A a été exprimé en fusion à la GFP dans les cellules MEFs  $p53^{-/-}$  et dans les cellules MEFs  $p53^{-/-}$   $dnmt1^{-/-}$  dépourvues de méthylation. BAZ2A est localisé aux chromocentres à la fois dans les cellules présentant un génome méthylé que dans les cellules ayant une perte globale de la méthylation du génome.

Les résultats sont présentés figure 23. On voit tout d'abord que le contrôle positif ZBTB4 colocalise avec les chromocentres, ainsi qu'il est publié (Filion et al. 2006). La GFP, elle, est diffuse dans l'ensemble de la cellule, et ne superpose pas avec les chromocentres. Zfp354b et ZNF354A (donnée non présentée) forment des foci nucléaires non superposés aux chromocentres, et qui sont sans doute les nucléoles. SOX14 montre un patron nucléaire diffus. Enfin ZCCHC4 est observé dans le cytoplasme et le noyau. Notre deuxième contrôle, Kaiso, ne colocalise pas avec les chromocentres, ainsi qu'il a déjà été décrit. Nous reviendrons sur ce phénomène dans la discussion.

Un candidat s'est avéré positif : il s'agit de BAZ2A qui superpose très clairement avec les chromocentres (figure 24). Nous avons donc réalisé une expérience contrôle dans un contexte où la méthylation du génome est perdue. L'expérience montre que la colocalisation entre BAZ2A et les régions centromériques persiste dans les cellules MEFs p53<sup>-/-</sup> dnmt1<sup>-/-</sup>. Cela signifie que la présence de cette protéine au centromère est sans doute indépendante de la méthylation de l'ADN.

BAZ2a contient un MBD, domaine connu pour lier l'ADN méthylé, et nous avons donc été surpris par ces premiers résultats. Pour mieux les comprendre, nous avons cherché à identifier le domaine minimum requis pour la localisation au niveau des régions centromériques. La figure 25 montre la transfection dans les cellules murines MEFs de différents domaines de BAZ2A fusionnés à la GFP. Le domaine DDT est localisé principalement dans le noyau mais il ne suffit pas à assurer la localisation aux chromocentres. En revanche, les motifs A-T Hook associés au domaine MBD permettent une superposition avec les chromocentres, bien qu'elle soit moins forte que celle obtenue avec la protéine de pleine longueur. Il est concevable d'imaginer que les régions flanquantes au MBD et aux motifs A-T Hook soient requises pour recruter BAZ2A aux chromocentres avec une affinité importante.

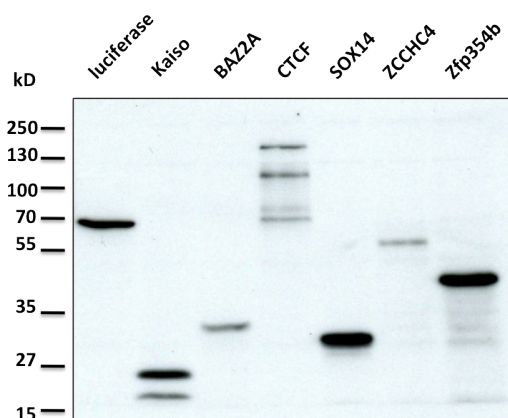


**Figure 25. Les motifs A-T hook associés au MBD permettent la localisation aux chromocentres.**

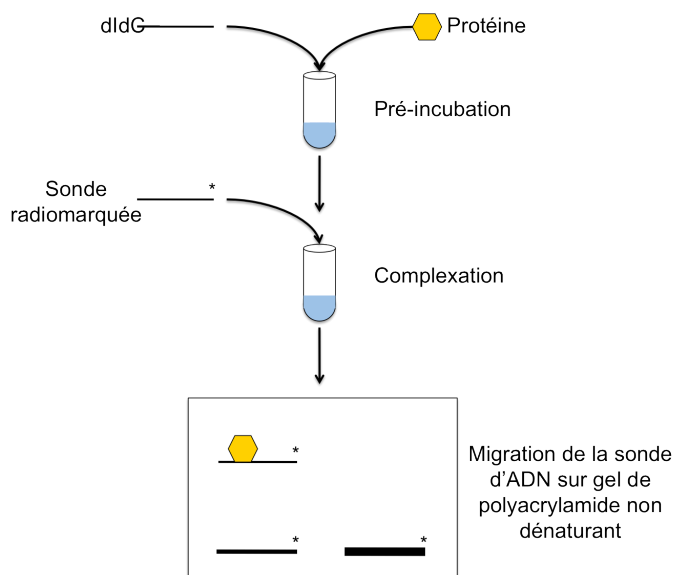
Les différents domaines de BAZ2A ont été exprimés en fusion à la GFP dans les cellules MEFs  $p53^{-/-}$ . BAZ2A est localisé aux chromocentres uniquement lorsque le domaine MBD est associé aux motifs A-T hook.

### 3.2.3 Résultats des expériences de retard sur gel

Les protéines ont été exprimées *in vitro* en lysat de réticulocytes de lapin. Cette transcription/traduction *in vitro* à l'avantage d'être un système eucaryote, et elle est bien plus rapide que l'expression en cellules bactériennes. En revanche, la quantité de protéines produites est inférieure au système bactérien. Le kit utilisé permet l'incorporation de lysine biotinylée dans la protéine naissante durant la traduction. La visualisation des protéines biotinylées est réalisée par western blot en utilisant la streptavidine-HRP (HorseRadish Peroxidase) suivi par une détection chemiluminescente (figure 26). Un plasmide exprimant la luciférase sert de contrôle positif à la transcription/traduction *in vitro*. Seul le domaine MBD de BAZ2A a été produit car, pour des raisons intrinsèques à la séquence d'ARN ou à sa structure, la synthèse *in vitro* des domaines MBD et DDT n'a pas été possible. Egalement, du au poids moléculaire trop important de la protéines Zfp354b, seuls les 13 doigts de zinc ont été synthétisés. De façon surprenante, la synthèse de son homologue humain : ZNF354A n'a pas pu être réalisée. Enfin, les protéines de pleine longueur CTCF, SOX14 et ZCCHC4 ont été produites. Nous avons utilisé la protéine Kaiso comme contrôle positif dans l'expérience de retard sur gel.

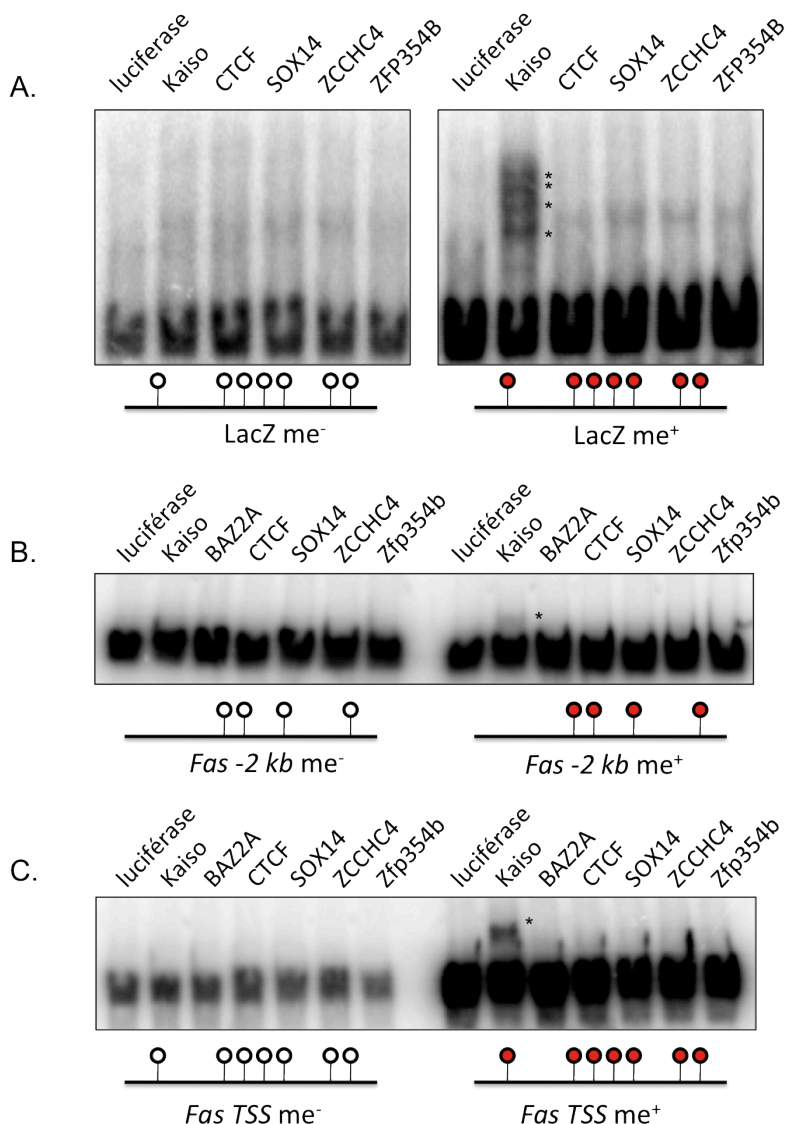


**Figure 26. Production *in vitro* des protéines candidates pour l'expérience de retard sur gel.** Les migrations observées correspondent au poids attendu pour les différentes protéines.



**Figure 27. Principe du retard sur gel.**

Les protéines sont pré-incubées avec un polymère dI-dC afin d'éliminer les interactions non spécifiques. La sonde radiomarquée est alors complexée avec les protéines et l'interaction est révélée par électrophorèse sur gel de polyacrylamide non dénaturant.



**Figure 28. Les protéines candidates ne se lient pas à l'ADN méthylé *in vitro*.**

Les sondes LacZ (A), *Fas* -2 kb (B) et *Fas* TSS (C) ont été incubées avec les différentes protéines candidates puis la migration des complexes a été réalisée par électrophorèse en gel d'acrylamide non dénaturant. La luciférase utilisée comme contrôle négatif est incapable de lier l'ADN méthylé ou non. Kaiso lie spécifiquement les trois sondes lorsqu'elles sont méthylées. Les protéines candidates n'interagissent ni avec les sondes méthylées ni avec les sondes non méthylées.

\* Complexe ADN/Kaiso.

Afin de tester par retard sur gel l'affinité des

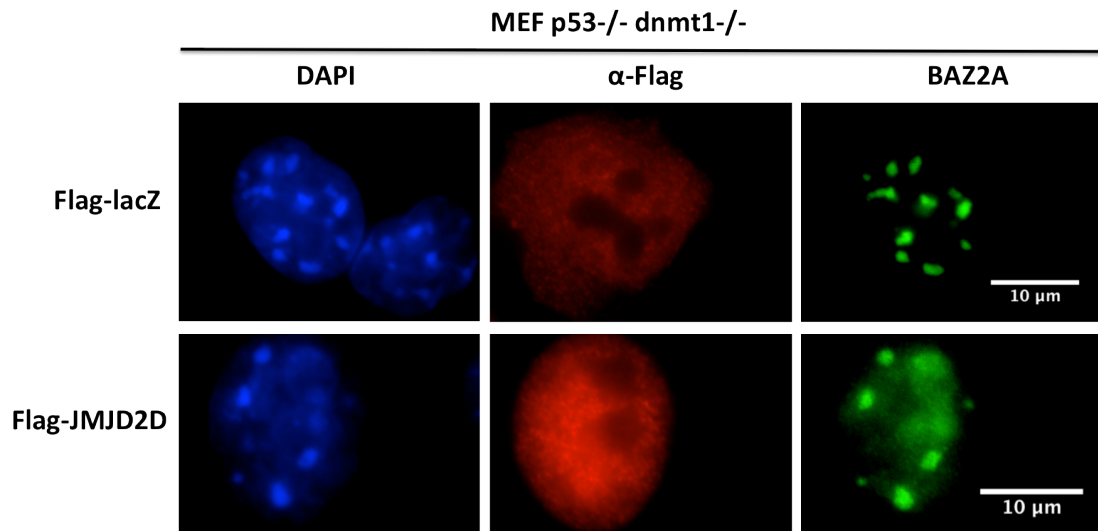
protéines pour l'ADN méthylé, trois sondes méthylées ont été utilisées :

- Une sonde correspond à la région - 2 kb du promoteur *Fas*. Le fragment d'une longueur totale de 307 paires de bases a été obtenu par PCR. Gazin et collègues ont montré par immunoprécipitation de chromatine que ZFP354B est présent sur cette région lorsqu'elle est méthylée.
- Une sonde correspond au site d'initiation de la transcription du promoteur *Fas*. Le fragment d'une longueur totale de 256 paires de bases a été obtenu par PCR. Gazin et collègues ont montré par immunoprécipitation de chromatine que CTCF est présent sur cette région lorsqu'elle est méthylée.
- Une dernière sonde correspond à un fragment du gène *lacZ*. Ce fragment d'une longueur totale de 161 paires de bases a été obtenu par PCR. Cette sonde n'a pas de relevance biologique, mais elle est un contrôle positif *in vitro*, connu pour lier Kaiso (Filion et al. 2006). Sa forte densité en CpGs laisse imaginer qu'elle pourrait lier diverses protéines reconnaissant l'ADN méthylé.

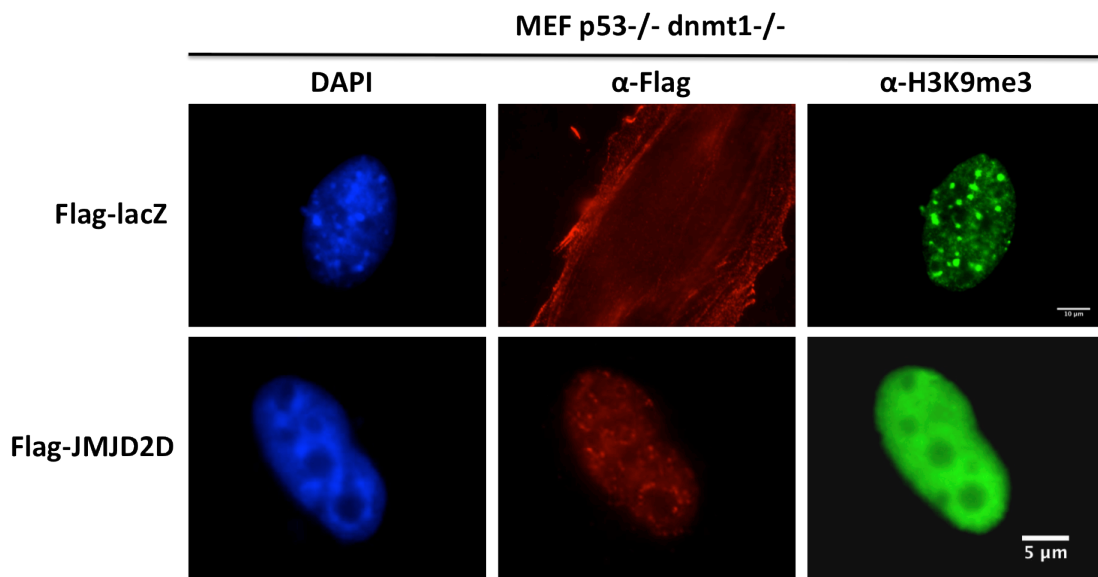
Les sondes nucléiques décrites ont été méthylées *in vitro* grâce à la méthylase bactérienne SssI, qui cible les dinucléotides CpG. La méthylation des sondes ont été vérifiée par digestion avec une enzyme sensible à la méthylation HpaII. La sonde méthylée, ainsi que la sonde non méthylée qui sert de contrôle négatif, sont ensuite radiomarquées au 32P.

Ensuite, l'affinité des protéines candidates pour l'ADN méthylé a été testée par retard sur gel (figure 27) avec les trois types de sondes cités précédemment. Comme attendu, la luciférase n'interagit pas avec l'ADN (figure 28). D'autre part, le contrôle positif Kaiso provoque bien un retard de migration des sondes méthylées *lacZ*, *Fas -2 kb* et *Fas TSS*, mais jamais des sondes non méthylées. Kaiso lie donc spécifiquement l'ADN méthylé. Par ailleurs, nous remarquons qu'aucune des protéines candidates ne se lie aux sondes méthylées, ni aux sondes non méthylées. Les protéines ne semblent donc pas affines pour l'ADN méthylé dans ces conditions.





**Figure 29. BAZ2A colocalise aux chromocentres en absence de H3K9me3.**  
 En bleu : contre coloration du noyau au DAPI, en rouge : immunofluorescence dirigée contre le tag flag permettant de contrôler la transfection, en vert : localisation de la GFP. BAZ2A colocalise avec les chromocentres à la fois dans les cellules transfectées par lacZ ou par JMJD2D malgré la perte de H3K9me3.



**Figure 30. Les cellules transfectées par JMJD2D montre une perte de la marque H3K9me3.**  
 En bleu : contre coloration du noyau au DAPI, en rouge : immunofluorescence dirigée contre le tag flag permettant de contrôler la transfection, en vert : immunofluorescence dirigée contre la marque H3K9me3. Les cellules transfectées par lacZ présentent un profil sauvage de H3K9me3 tandis que les cellules transfectées par JMJD2D montrent une diffusion de H3K9me3 dans le noyau.

**Table 14. Résumé des résultats.**

	Localisation aux chromocentres		Fixation à la sonde me <sup>+</sup> <i>in vitro</i>		
	En cellules Dnmt1 <sup>+/+</sup>	En cellules Dnmt1 <sup>-/-</sup>	LacZ	<i>Fas</i> -2 kb	<i>Fas</i> TSS
<b>ZBTB4 (+)</b>	+	-	+	ND	ND
<b>GFP (-)</b>	-	-	ND	ND	ND
<b>Kaiso (+)</b>	-	-	+	+	+
<b>SOX14</b>	-	-	-	-	-
<b>BAZ2A</b>	+	+	-	-	-
<b>ZCCHC4</b>	-	-	-	-	-
<b>ZNF354A</b>	-	-	-	-	-
<b>Zfp354B</b>	-	-	-	-	-
<b>CTCF</b>	ND	ND	-	-	-

\* Donnée non présentée.

### 3.3 Discussion

Par une approche basée sur la littérature, des protéines candidates ont été sélectionnées. Ces protéines ont alors été analysées *in vitro* et *in vivo* afin de savoir si elles sont capables de lier l'ADN méthylé. Les résultats apparaissent négatifs, mais les précisions suivantes s'imposent.

#### 3.3.1 Limites à l'interprétations des résultats

Parmi les protéines testées, ZCCHC4 n'est pas nucléaire, et il est donc très peu vraisemblable qu'elle lie l'ADN méthylé. Pour les autres, il faut souligner qu'un manque de colocalisation avec les chromocentres ne permet pas de conclure qu'une protéine ne lie pas l'ADN méthylé. En effet, certaines protéines, dont Kaiso, lient clairement l'ADN méthylé *in vitro* et *in vivo*, mais ne colocalisent pas avec les chromocentres. Ceci s'explique par le fait qu'elles se lient à une séquence méthylée spécifique qui est absente des chromocentres. SOX14, qui présente un patron nucléaire diffus, pourrait être dans le même cas de figure que Kaiso. Quant à ZFP354b, qui semble concentré aux nucléoles, elle pourrait éventuellement lier une séquence de l'ADN ribosomique, qui est connu pour être partiellement méthylé.

On voit que les expériences de localisation sont, par nature, seulement

indicatives, et non conclusives. Les expériences de retard sur gel sont donc déterminantes pour renforcer les indications fournies par l'observation *in vivo*. Cependant elles présentent elles aussi diverses limitations. La principale est que l'affinité des protéines pour l'ADN méthylé y est testée pour une séquence nucléique spécifique et définie par l'expérimentateur. Comme dans l'exemple de Kaiso, outre la nécessité des CpG méthylés, les protéines pourraient requérir des séquences flanquantes spécifiques afin d'établir une interaction avec l'ADN méthylé. Il est donc envisageable que la sonde utilisée ne contient pas les séquences cibles, et livre un résultat faussement négatif. Pour diminuer cette probabilité, nous avons utilisé trois sondes différentes, dont deux correspondent à des cibles connues *in vivo*. Les résultats obtenus avec ces trois sondes sont négatifs. Bien qu'ils ne permettent pas à eux seuls de conclure, en combinaison avec les expériences *in vivo*, ils suggèrent de façon forte que les candidats ne lient pas l'ADN méthylé dans les conditions utilisées.

Plusieurs choses restent à éclaircir. Gazin et collègues ont montré par immuno-précipitation de chromatine que CTCF et ZFP354B lient le promoteur Fas méthylé, mais pas le promoteur non méthylé. Comment l'expliquer si les protéines ne lient pas l'ADN méthylé? Il faut imaginer que les protéines reconnaissent le patron de chromatine qui accompagne la méthylation, ou qu'elles lient le promoteur de façon indirecte, en reconnaissant une autre protéine qui reconnaît l'ADN méthylé et/ou la chromatine répressive. Quant aux autres candidats, pourquoi sont-ils nécessaires au maintien de la méthylation du promoteur Fas s'ils ne s'y lient pas? Il est possible qu'ils agissent de façon indirecte, en régulant l'expression d'un gène qui joue sur la méthylation du promoteur.

### 3.3.2 Interprétations des résultats de colocalisation de BAZ2A

Les résultats de localisation GFP de BAZ2A montrent que la protéine colocalise avec les chromocentres. Cependant, cette superposition est indépendante de la méthylation de l'ADN. Une possibilité serait que BAZ2A est recruté aux chromocentres via la méthylation des histones. En effet, la méthylation de la lysine 9 de l'histone H3 est associée à la répression transcriptionnelle. De plus, les régions génomiques enrichies en ADN

méthylé sont corrélées avec la diméthylation de H3K9, et il est possible d'un point de vue structural que les domaines protéiques contenus dans BAZ2A reconnaissent les histones modifiées.

Afin de tester cette hypothèse, nous avons voulu savoir si BAZ2A colocalise avec les chromocentres lorsque la di- et triméthylation de la lysine 9 de l'histone H3 sont supprimées par expression de la déméthylase JMJD2D. La co-transfection de la déméthylase JMJD2D avec BAZ2A a donc été réalisée dans des cellules de souris MEFs p53<sup>-/-</sup> Dnmt1<sup>-/-</sup>. Une immunofluorescence dirigée contre le tag permet de visualiser en rouge les cellules transfectées par flag-lacZ ou par flag-JMJD2D. La localisation de BAZ2A est quant à elle observée en vert par fusion à la GFP. Comme le montre la figure 29, ni la co-transfection du contrôle lacZ ni la co-transfection de JMJD2D avec BAZ2A n'induisent la relocalisation de BAZ2A hors des chromocentres dans un contexte d'absence de méthylation de l'ADN. Nous avons vérifié par ailleurs que JMJD2D était bien active (Figure 30). Deux interprétations sont possibles pour expliquer le fait que BAZ2A soit encore recruté aux chromocentres en absence de méthylation de l'ADN et de modification de H3K9. Premièrement, la déplétion de ces marques dans les cellules utilisées est forte mais pas totale. Il est possible que le niveau résiduel suffise à recruter BAZ2A, bien qu'il ne suffise pas à attirer ZBTB4 (données non présentées). Une autre possibilité serait que le recrutement de BAZ2A ne dépende ni de la méthylation de l'ADN, ni de la modification des histones, mais d'un autre élément. Il pourrait s'agir d'un ARN non-codant. En effet, il a été montré récemment que BAZ2A est capable d'interagir avec un ARN non codant afin de cibler le recrutement du complexe NoRC à l'ADN ribosomique (Schmitz et al. 2010). BAZ2A est également nécessaire à l'établissement de l'hétérochromatine péricentrique (Guetg et al. 2010). Or il s'avère que les satellites des régions centromériques peuvent être transcrites permettant probablement la régulation de l'assemblage de l'hétérochromatine chez les mammifères (Bouzinba-Ségard et al. 2006). Il serait donc intéressant de vérifier si le recrutement de BAZ2A aux chromocentres dépend de transcrits péricentriques.

### 3.3.3 Perspectives

Enfin, il a récemment été montré qu'une protéine de la famille des ZBTB est associée avec le syndrome ICF de type 2 (de Greef et al. 2011). Des mutations de cette protéine nommée ZBTB24 provoque une hypométhylation des séquences centromériques. ZBTB24 s'avère donc être un candidat potentiel pour lier l'ADN méthylé et permettre le maintien ou la mise en place de la méthylation de l'ADN sur les régions centromériques.

### 3.4 Matériels et méthodes

#### Séquences des oligonucléotides utilisés

Les séquences des oligonucléotides sont listées dans la table 16. Les oligonucléotides ont été synthétisés par *Sigma-Aldrich*.

**Table 15. séquence des oligonucléotides utilisés pour les clonages.** (les sites de restrictions en 5' sont indiqués en rouge et les sites de restrictions en 3' sont indiqués en bleu).

Oligo	Gène	Séquence	Commentaire
MJ1	SOX14	GGGG <b>GAATTC</b> ATGTCCAAACCTTCAGACC	Amorce 5', site EcoRI
MJ2	SOX14	GGGG <b>GGATCC</b> TTACATGGCCGTAGCGT	Amorce 3', site BamHI
MJ5	BAZ2A	GGGG <b>GAATTC</b> GTTTGTCCCGCAGCTTCTAC	Amorce 5', site EcoRI
MJ6	BAZ2A	GGGG <b>GGATCC</b> TCTACTTCAGACCGCCAGT	Amorce 3', site BamHI
MJ7	ZCCHC4	GGGG <b>GAATTC</b> GCCTCCAGGAATGGGTTT	Amorce 5', site EcoRI
MJ8	ZCCHC4	GGGG <b>AGATCT</b> ATGGGCTCTTTCCCTCCTT	Amorce 3', site BglII
MJ21	ZNF354A	GGGG <b>GTCGAC</b> AAGGAAACAGCCTCAAACAG	Amorce 5', site Sall
MJ22	ZNF354A	GGGG <b>GGATCC</b> GGGGTCTCTTCGATATGAA	Amorce 3', site BamHI
MJ23	Zfp354b	GGGG <b>GTCGAC</b> AGCAAAAGATTGCAGGAGAA	Amorce 5', site Sall
MJ24	Zfp354b	GGGG <b>AGATCT</b> CTCATGCAAATCTATGTTCA	Amorce 3', site BglII
MJ35	BAZ2A	GGGG <b>GGATCC</b> CGTTTGCCAGTAATTGCCTG	Amorce 3', site BamHI

#### Clonage moléculaire

Les clones d'ADNc utilisés pour amplifier les gènes candidats ont été obtenus via la collection publique *ImaGenes* ([www.imagenes-bio.de](http://www.imagenes-bio.de)). *ZCCHC4* n'étant pas disponible dans la banque, le gène a été obtenu par PCR à partir d'une préparation d'ADNc de cellules humaines HEK293 (*BD Biosciences*).

La polymérase Phusion (Finnzymes) est utilisée pour réaliser les amplifications par PCR des clones. Les produits d'amplification sont ensuite purifiés par passage sur mini colonne *NucleoSpin Extract II* (Macherey-Nagel) et repris dans 50 µl de tampon d'élution AE (5mM Tris/HCl, pH 8.5). 1 µg de produit purifié est digéré par les enzymes choisies (**table 17**) pendant 2 heures à 37°C puis purifié sur mini colonne *NucleoSpin Extract II* et repris dans 30 µl de tampon d'élution AE. Les fusions Green Fluorescent Protein (GFP) sont obtenues en clonant les produits digérés dans pEGFP-C2 (Clontech). 2 µg de pEGFP-C2 sont digérés pendant 3 heures à 37°C puis purifiés par mini colonne *NucleoSpin Extract II* et repris dans 30 µl de tampon d'élution AE. 50 ng de vecteur sont combinés avec 3 à 6 fois d'excès molaire d'insert puis le volume est ajusté à 10 µl avec du TE (10 mM Tris, pH 8 et 1 mM EDTA). 10 µl de tampon 2X Quick Ligation Buffer (*New England Biolabs*, 132 mM Tris-HCl, 20 mM MgCl<sub>2</sub>, 2 mM dithiothreitol, 2 mM ATP, 15% Polyethylene glycol pH 7.6) sont ajoutés ainsi que 1 µl de Quick T4 DNA Ligase (*New England Biolabs*). La réaction de ligation est incubée 5 minutes à température ambiante. Pour la transcription/ traduction *in vitro*, les ADNc ont été clonés dans pCITE-4a (*Novagen*).

**Table 16. récapitulatif des clonages réalisés.** (Hs : *Homo sapiens*. Mm : *Mus musculus*)

Gène	Source de l'ADNc amplifié	Digestion du produit de PCR	Sites d'insertion dans pEGFP	Sites d'insertion dans pCITE 4a
BAZ2A (Hs)	Clone Imagenes DKFZp781B109Q	EcoRI/BamHI	EcoRI/BamHI	EcoRI/BglII
SOX14 (Hs)	Clone Imagenes IRAMP995O138Q	EcoRI/BamHI	EcoRI/BamHI	EcoRI/BglII
ZNF354A (Hs)	Clone Imagenes IRATp970A0563D	Sall/BamHI	Sall/BamHI	Sall/BglII
Zfp354b (Mm)	Clone Imagenes IRAMP995O1916Q	Sall/BamHI	Sall/BamHI	Sall/BglII
ZCCHC4 (Hs)	ADNc de cellules HEK293	EcoRI/BglII	EcoRI/BamHI	EcoRI/BglII

### Crible PCR sur colonies bactériennes

Lors des différents clonages, les colonies bactériennes sont testées par PCR avec une paire d'oligonucléotides appropriée pour déterminer si elles contiennent ou non un insert. Pour cela, 12.5 µl de 2X BioMix™ Red

(*Bioline*, Taq DNA polymerase, MgCl<sub>2</sub>, dNTPs) sont ajoutés à 2.5 µl de chaque oligonucléotide 10 mM et le volume est ajusté à 25 µl avec de l'eau. Un cône est alors utilisé pour piquer une colonie pour être ensuite plongé dans le tube de PCR. Le programme PCR mis au point est le suivant : 94°C 3 min., 94°C 30 sec., 55°C 30 sec., 72°C 2 min., 24 cycles, 72°C 7 min. et maintien à 4°C. Les résultats sont alors analysés par électrophorèse sur gel d'agarose 1.5%.

### **Transfection de cellules de mammifères**

Les cellules de souris MEFs (Mouse Embryonic Fibroblast) p53<sup>-/-</sup> et MEFs p53<sup>-/-</sup>/Dnmt1<sup>-/-</sup> sont cultivées en DMEM (*Gibco*) 10% SVF (sérum de veau foetal) sans antibiotique à 37°C. La veille des transfections, les cellules sont trypsinées etensemencées à 50 000 cellules par puits dans des plaques de culture à 24 puits. Au moment des transfections, les cellules sont alors à 80- 85% confluentes. La transfection est réalisée par la Lipofectamine 2000 (*Invitrogen*). Chaque puit est transfecté par 800 ng de plasmide et 1.8 µl de Lipofectamine 2000 dilués dans 100 µl de milieu Opti-MEM (*Gibco*). 4 à 6 heures après la transfection le milieu est remplacé par du DMEM 10% SVF sans antibiotique. Les cellules sont observées après 21 heures de transfection.

### **Fixation des cellules exprimant la GFP et microscopie**

Pour observation de la GFP et contre-coloration du noyau au DAPI, les cellules sont lavées dans du Phosphate-Buffered Saline (PBS), fixées pendant 10 minutes à température ambiante avec une solution de paraformaldéhyde (PFA) à 2% dans le PBS, lavées à nouveau avec du PBS, puis perméabilisées 5 minutes à 4°C dans du PBS-0.5% Triton X-100. Après un nouveau rinçage au PBS, les cellules sont incubées 3 minutes avec 0.3 µg/ml de 4',6'-diamidino-2-phenylindole (DAPI) dans du PBS. Les lamelles sont montées dans du Vectashield H-1000 (*Vector Laboratories Inc.*) et les images sont acquises avec un microscope Zeiss AxioImager Z1.

### **Immunofluorescence et microscopie**

Dix-huit heures après transfection, les cellules sont lavées dans du PBS, fixées pendant 10 minutes à température ambiante avec de la PFA 2%, lavées à nouveau avec du PBS puis perméabilisées 5 minutes à 4°C dans du

PBS-0.5% Triton X-100. Après un nouveau rinçage au PBS, les sites interagissant de façon non spécifique avec les protéines sont saturés par une incubation de 30 minutes à température ambiante avec 1% d'albumine sérique bovine (BSA) dans du PBS. L'anticorps primaire est dilué dans 1% BSA/PBS et 15 µl sont déposés sur chaque lamelle pendant 1 heure à température ambiante. Les cellules sont lavées 35 minutes dans du PBS-Tween20 0.1% par bains de 5 minutes. L'anticorps secondaire dilué au 1/1000 est alors incubé dans le noir avec 1% BSA/PBS pendant 30 minutes. Après 25 minutes de rinçage dans du PBS-Tween20 0.1%, les cellules sont colorées au DAPI (0.3 µg/ml) puis lavées et montées dans du Vectashield H-1000 (*Vector Laboratories Inc.*).

### **Transcription/traduction *in vitro***

La transcription/traduction *in vitro* est réalisée avec le kit TNT Quick Coupled Transcription/Translation Systems (*Promega*). 1 µg de construction pCITE-4a est ajouté à 20 µl de TNT Quick Master Mix (lysate de réticulocyte de lapin, tampon de réaction, ARN polymérase, mélange d'acides aminés et RNasin : inhibiteur de ribonucléase) avec 0.5 µl de méthionine 1 mM et 1 µl de Transcend Biotin-lysyl-tRNA (qui permet que les protéines synthétisées contiennent des lysines couplées à la biotine en position epsilon). Le volume est ajusté à 25 µl avec de l'eau et la réaction est incubée 90 minutes à 30°C. Les produits de réaction sont analysés par électrophorèse sur gel SDS-PAGE et Western blot.

### **Analyse par Western blot des produits de traduction**

Un µl de réaction de traduction est ajouté à 15 µl de tampon de charge contenant du SDS puis dénaturé à 90°C pendant 5 minutes. Après refroidissement, les échantillons sont chargés sur un gel de polyacrylamide 10% et soumis à électrophorèse à 100V jusqu'à la sortie du bleu de bromophénol. Après migration, le gel est équilibré dans du tampon de transfert (192 mM glycine, 25 mM Tris-base et 20% éthanol) puis transféré pendant 1 heure à 100V sur une membrane PVDF (polyfluorure de vinylidène) préalablement activée par immersion dans l'éthanol. La détection des produits est alors réalisée à l'aide du kit Transcend Chemiluminescent Substrate (*Promega*). La membrane est bloquée pendant 1



heure dans du PBST (PBS + 0.5% Tween20). Puis la membrane est incubée pendant 60 minutes sous agitation avec une dilution 1/10 000 de Streptavidin-HRP/PBST. Après rinçage en PBST, la membrane est incubée pendant 1 minute avec un mélange de substrat chimioluminescent (Transcend Chemiluminescent Substrate A et B). La membrane est alors exposée sur un film autoradiographique (Amersham Hyperfilm MP, GE Healthcare).

### **Amplification et méthylation *in vitro* des sondes nucléiques**

Les trois sondes utilisées sont décrites dans la table 18. Pour méthyler *in vitro* les dinucléotides CG de ces sondes, nous avons procédé comme suit. Les réactions sont réalisées dans un volume final de 100 µl avec 2 µg d'ADN, 10 µl de tampon NEB 2 (10 mM Tris-HCl pH 7.9, 10 mM MgCl<sub>2</sub>, 50 mM NaCl, 1 mM dithiothreitol à 25°C, *New England Biolabs*), 1 µl SAM (S-adenosylmethionine, *New England Biolabs*) et 4 unités de SssI (*New England Biolabs*). La réaction est placée 6 heures à 37°C mais après 3 heures d'incubation, 1 µl de SAM est ajouté car ce cofacteur est instable à 37°C. La sonde méthylée est ensuite purifiée par passage sur mini colonne *NucleoSpin Extract II* (*Macherey-Nagel*) et reprise dans 30 µl de tampon d'élution AE. Le degré de méthylation de la sonde peut être vérifié par digestion avec une enzyme de restriction sensible à la méthylation: HpaII. Après 1 heure à 37°C, les produits de digestion sont analysés sur gel d'agarose 1.5% (figure 31).

La sonde a été incubée avec l'enzyme HpaII, qui coupe les sites C/CGG. La méthylation préalable de ces sites empêche l'enzyme d'agir. On constate que la sonde méthylée est partiellement protégée contre la digestion.

Les sondes nucléiques sont marquées au <sup>32</sup>P par la polynucléotide kinase du phage T4 (PNK, *New England Biolabs*). La réaction est effectuée dans un volume de 50 µl avec 5 µl de tampon 10X PNK (70 mM Tris-HCl (pH 7.6), 10 mM MgCl<sub>2</sub>, 5 mM dithiotreitol, *New England Biolabs*), 100 à 200 ng d'ADN, 20 unités de PNK et 0.75 MBq de [μ-<sup>32</sup>P]-ATP (*MP Biomedicals*). Les tubes sont placés à 37°C pendant 45 minutes puis les sondes radiomarquées sont purifiées par colonnes G-25 Sephadex (Quick Spin Columns, *Roche*).

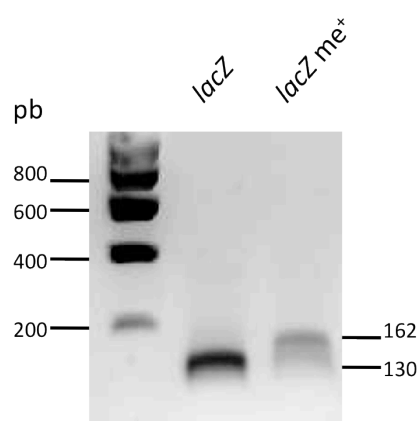
### Analyse de la liaison à l'ADN méthylé par retard sur gel

Les protéines produites *in vitro* en lysat de réticulocytes sont pré-incubées avec 1 µl de polymère dI-dC (1 µg/µl) pour saturer toute interaction non spécifique protéine-ADN. La réaction s'effectue à 4°C pendant 15 minutes dans un volume final de 20 µl contenant 10 µl de tampon d'assemblage 2X (50 mM HEPES pH 7.5, 100 mM KCl, 5mM MgCl<sub>2</sub>, 2 µM ZnSO<sub>4</sub>, 0.2% NP-40, 10% glycérol et 2 mM dithiotreitol), et 1 à 3 µl de protéines. Puis 2 µl de sondes radiomarquées sont complexés avec les protéines pendant 15 minutes à 4°C. Les complexes sont ensuite déposés sur un gel de TBE/Acrylamide 5% et migrent 1 heure et demi à 50V à 4°C. Le gel est alors séché 50 minutes à 72°C sur papier Whatman et analysé par Phosphoimager (*Molecular Dynamics*).

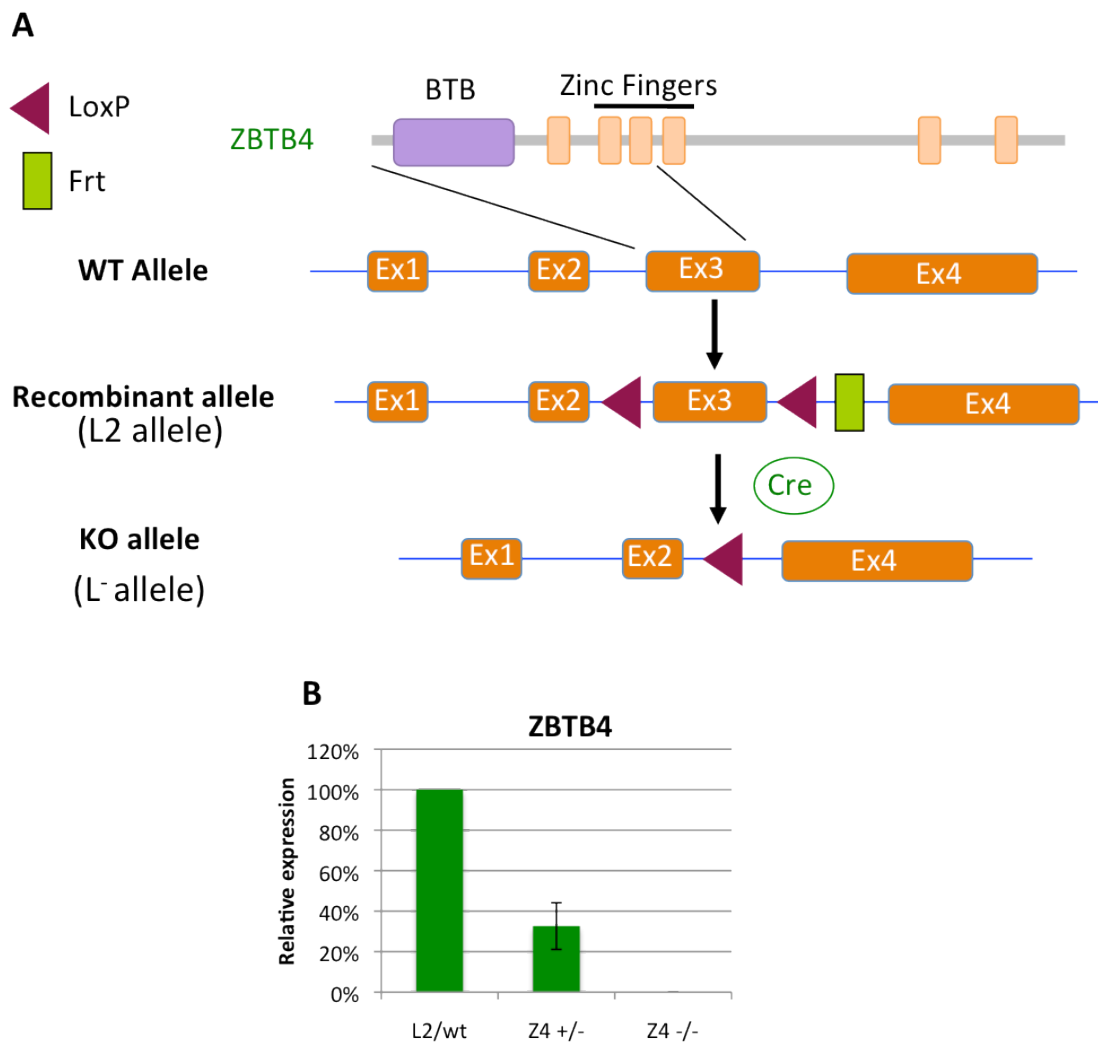
**Table 17. liste des sondes utilisées pour les expériences de retard sur gel.**

Sonde	Taille (pb)	Nombre de CG	Amorce 5'	Amorce 3'
LacZ	162	22	GF048 : TAATGCGATTAGTGAATTCGCGCGCT GGCTACCGCGCAT	GF050 : TCATTATGTGATAATGCCGCGCG TCGTGATTAGCGCCGT
Fas -2 kb	307	6	MJ19 : GCGTTGCCATAGCATGAACT	MJ20 : GAGTTAGGGGACCATAGTCA
Fas TSS	256	11	MJ33 : GCCGCCTGTGCAGTGGTGA	MJ34 : CTGTGTGTGGGCAGCCTGCGGC

### Marquage des sondes pour retard sur gel



**Figure 31. Vérification de la méthylation de la sonde *lacZ*.**



**Figure 32. Stratégie et Confirmation du KO.**

A. L'exon 3, flanqué de deux sites LoxP est délété sous l'action de la Cre recombinase. B. Confirmation au niveau ARN par qPCR en utilisant des primers amplifiant une région chevauchant les exons 3 et 4.

## Partie III : Régulation des éléments répétés par ZBTB4

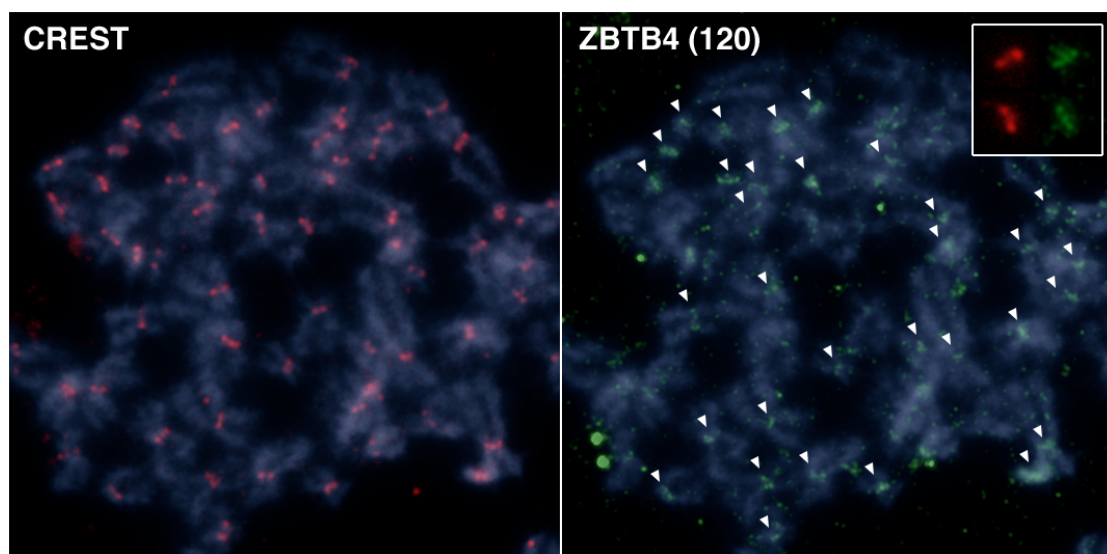
### Chapitre 1 : introduction

Afin de comprendre le rôle de Zbtb4, nous avons étudié l'implication de cette protéine dans la régulation des éléments répétés. Nous avons pu voir que la méthylation de l'ADN est indispensable au maintien de l'intégrité du génome via la répression des éléments transposables et permet la formation des structures hétérochromatiques au niveau des centromères (Partie I, Chapitre 3.4). Ces régions, fortement méthylées, s'avèrent être de potentielles cibles pour les MBPs. L'absence de MeCP2 dans le cerveau est d'ailleurs corrélée avec la réexpression des satellites majeurs (Skene et al. 2010).

Comme nous avons pu le voir, ZBTB4 a été identifiée et caractérisée par notre laboratoire pour sa capacité à lier l'ADN méthylé (Filion et al. 2006). En particulier, dans des fibroblastes embryonnaires murins (MEF), une protéine étiquetée GFP-ZBTB4 se localise au niveau des régions denses de DAPI, riches en ADN satellites majeurs et mineurs, composant les centromères (Filion et al. 2006). La perte de méthylation de ces ADN dans un contexte DNMT1-/- abolit le recrutement de GFP-ZBTB4 indiquant que GFP-ZBTB4 se lie à des motifs CpG méthylés au sein de ces ADN satellites (Filion et al. 2006). Cette localisation suggère que ZBTB4 pourrait être impliqué dans la régulation des séquences satellites et, probablement, d'autres éléments répétés du génome comme par exemple les transposons. Par ailleurs, nous avons montré que ZBTB4 endogène localise au niveau des centromères au cours de la mitose de cellules humaines (figure 33). Nous avons donc effectué des analyses d'expression de certains éléments répétés dans des contextes mutants pour *Zbtb4* à la fois par PCR semi-quantitative et par qPCR. Par ailleurs, nous avons testé si Zbtb4 se lie *in vivo* à ces mêmes éléments par des expériences d'immunoprécipitation de la chromatine (ChIP).

Afin de réaliser ces expériences *in vivo* et dans un contexte physiologique, l'obtention d'une souris délétée pour le gène *Zbtb4* a été

entreprise en collaboration avec l'Institut Clinique de la Souris (ICS) de Strasbourg. Les souris mutantes portent une délétion de l'exon 3 du gène *Zbtb4* contenant le codon ATG initiateur de la traduction et codant pour le domaine BTB/POZ ainsi que pour deux des trois doigts de zinc impliqués dans la reconnaissance à l'ADN méthylé (figure 32). La validation et la caractérisation des souris mutantes pour *Zbtb4* sont le sujet de recherche de Laetitia Kasprzyk, étudiante en thèse au sein du laboratoire. Elle a été en mesure de confirmer, au niveau ADN par Southern blot, puis au niveau ARN par PCR quantitative (qPCR), la délétion du gène *Zbtb4*. La confirmation au niveau protéine est actuellement en cours et s'avère compliquée de par l'absence d'un anticorps reconnaissant la protéine *Zbtb4* murine de manière efficace. Les souris *Zbtb4*<sup>-/-</sup> sont viables et l'étude de leur phénotype est en cours. Pour le moment, aucun défaut morphologique n'a été observé.



**Figure 33. Localisation de ZBTB4 dans des cellules humaines.** ZBTB4 localise au niveau de certains centromères au cours de la mitose. Le CREST permet de marquer les kinétochores. ZBTB4 est révélée par l'anticorps ZBTB4 #120.

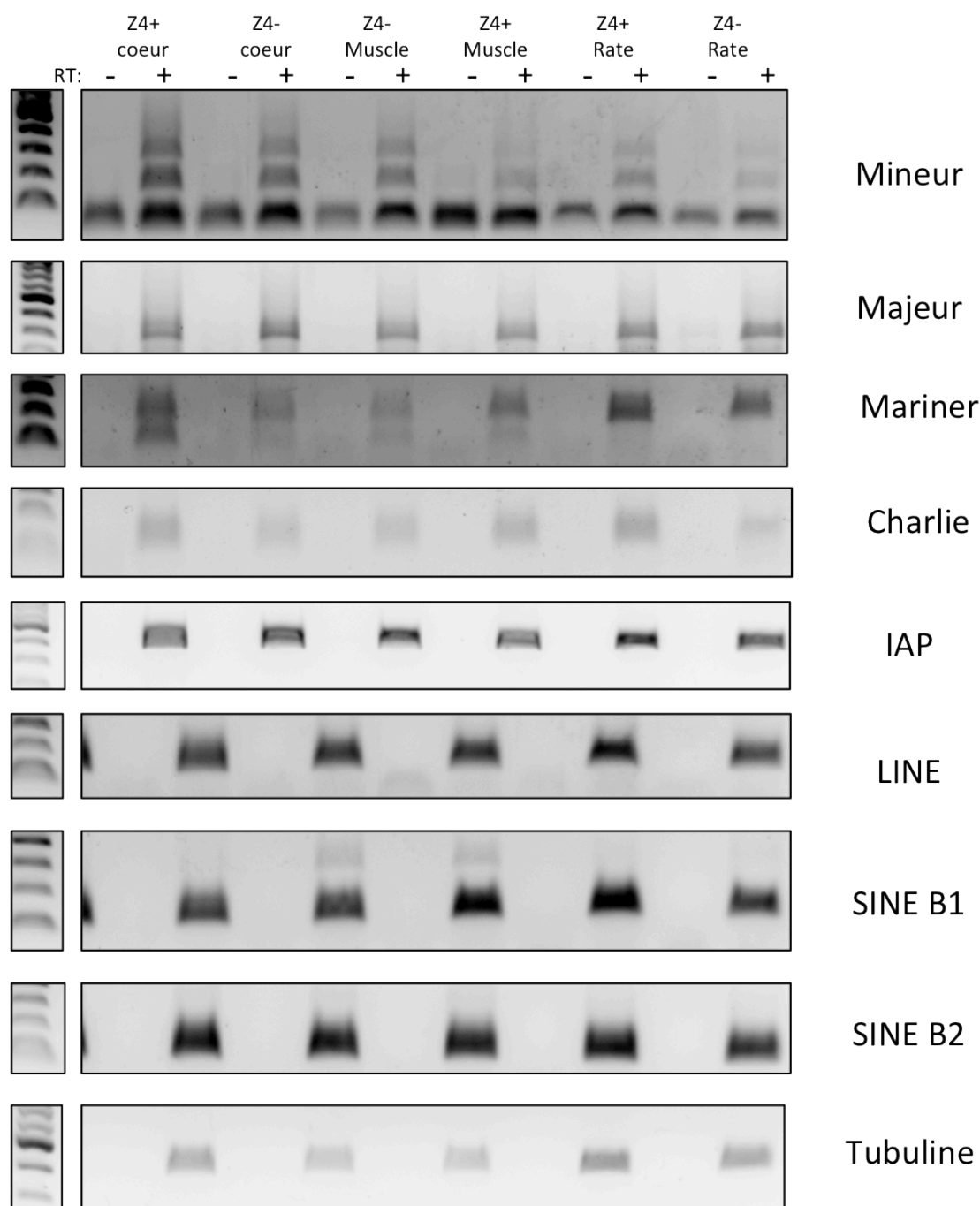
## Chapitre 2 : résultats

### 2.1 Analyse d'expression des éléments répétés

Dans un premier temps, nous avons étudié l'expression des satellites mineurs et majeurs présents, respectivement, aux niveaux des régions centromériques et péricentromériques ainsi que les éléments transposables Mariner, Charlie, IAP, LINE, SINE B1 et SINE B2. Pour cela, nous avons extrait les ARNs totaux de différents tissus chez les souris sauvages et mutantes pour *Zbtb4*. Les ARNs ont alors été rétrotranscrits en ADNc puis analysés soit par PCR semi-quantitative soit par qPCR. Les résultats sont présentés figure 34 et 35. Dans les souris délétées pour *Zbtb4*, d'après les résultats de la PCR semi-quantitative, aucun élément ne semble ré-exprimé quelque soit le tissu testé (ex : le cœur, le muscle et la rate). Cependant, nous avons réalisé des expériences de qPCR afin d'obtenir une plus grande précision de calcul, une standardisation des expériences ainsi qu'une meilleure sensibilité. Pour des raisons de quantité de matériel disponible, les expériences de qPCR ont été réalisées à partir de foie et de rein de souris sauvages ou mutantes pour *Zbtb4* (*Zbtb4*<sup>+/-</sup> et *Zbtb4*<sup>-/-</sup>). Dans le foie, seul les satellites mineurs montrent une surexpression de 2 à 5 fois à travers trois expériences indépendantes (figure 36 et annexe 1). Dans les reins, des résultats préliminaires semblent indiqués une réexpression des éléments Charlie, LINE et SINE d'environ deux fois. Cependant, ces résultats nécessitent d'être reproduit afin de confirmer la dérégulation de ces éléments dans les souris mutantes pour *Zbtb4*.

### 2.2 Localisation de ZBTB4 *in vivo*

Dans le but d'étudier la localisation de *Zbtb4 in vivo* et de déterminer si cette dernière peut réguler de façon directe les éléments répétés, nous avons effectué des expériences de ChIP.



**Figure 34. Analyse d'expression des éléments répétés par PCR semi-quantitative.**

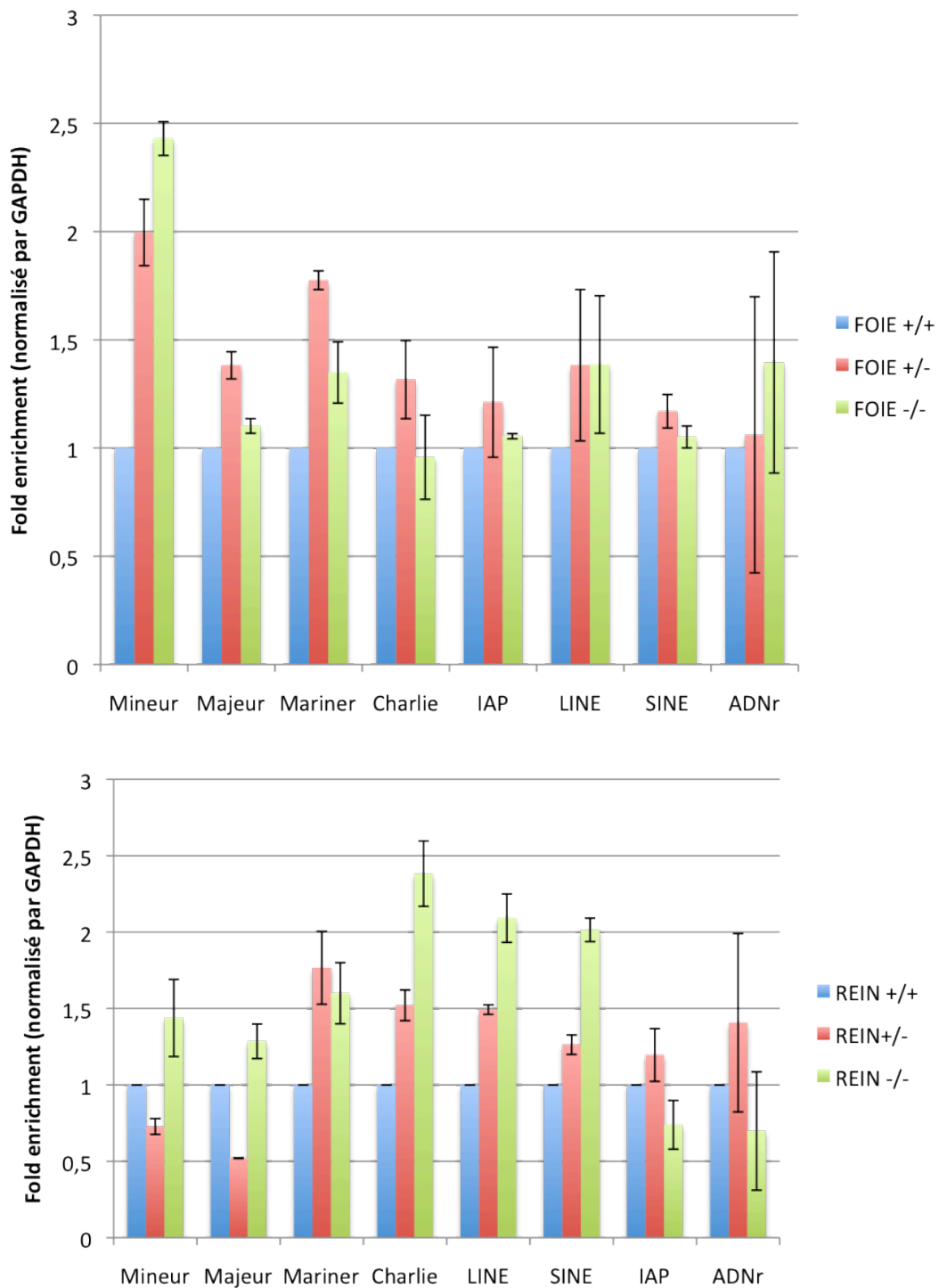
L'expression de différents éléments répétés a été analysée par PCR semi-quantitative dans le cœur, le muscle et la rate de souris sauvages et mutantes pour ZBTB4. L'expression de ces différents éléments répétés reste inchangée entre les tissus des souris sauvage et les souris délétées pour ZBTB4.

RT : rétrotranscription.

Dans un premier temps, nous avons testé la liaison de la protéine Zbtb4 endogène avec les satellites mineurs et majeurs, en utilisant l'anticorps #120 (figure 36). Nous pouvons constater par qPCR qu'il n'y a pas d'enrichissement des satellites mineurs et majeurs dans les immunoprécipitations de ZBTB4 en comparaison avec les contrôles négatifs que sont les billes seules et les immunoglobulines. En parallèle nous avons observé un enrichissement des satellites majeurs et mineurs dans des immunoprécipitations de la marque d'histone H3K9me3. Ce résultat valide la procédure de ChIP et la qualité de la préparation de chromatine puisque les satellites majeurs et mineurs sont associés à une structure chromatine répressive enrichie en H3K9me3 (Lehnertz et al. 2003 ; Martens et al. 2005). A ce stade, 2 explications sont possibles : 1- ZBTB4 endogène ne se lie pas aux ADN satellites ou 2- les anticorps utilisés ne permettent pas d'isoler les complexes ZBTB4/ADN.

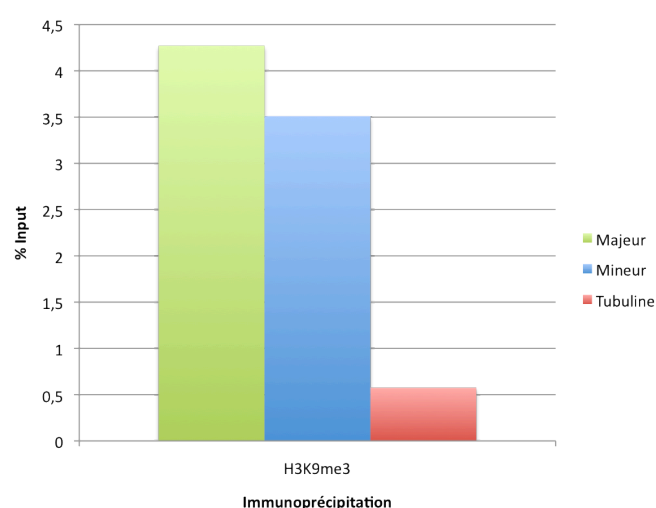
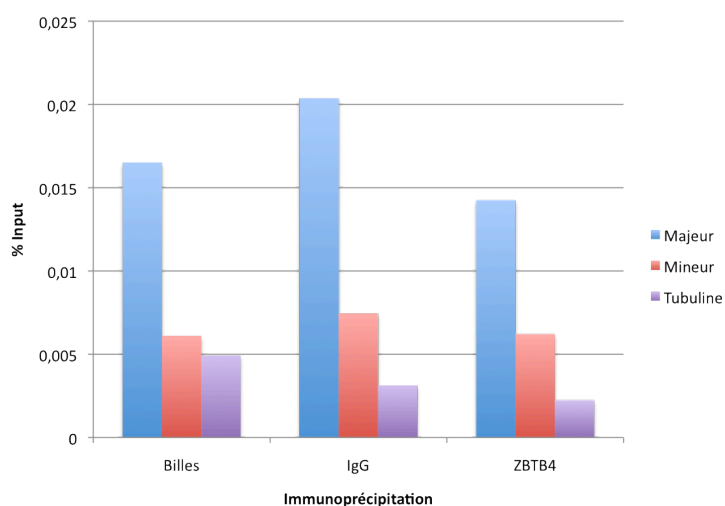
Nous avons donc décidé d'analyser la localisation de Zbtb4 dans un contexte de surexpression. Pour cela, nous avons transfecté un plasmide contenant ZBTB4 humaine fusionnée à la GFP et, en parallèle un plasmide GFP seul. Nous avons alors immunoprécipité GFP-ZBTB4 en utilisant un anticorps anti-GFP (figure 37). Nous avons observé que les ADN satellites majeurs et mineurs sont enrichis dans des immunoprécipitations de GFP-ZBTB4 comparé à des immunoprécipitations de GFP seul. Ces résultats suggèrent que ZBTB4 se lie aux séquences satellites majeures et mineures formant le centromère et pericentromère, ce qui corrèle avec les données d'immunofluorescence (Filion et al. 2006). Toutefois, l'enrichissement est faible et la tubuline, en tant que contrôle négatif, présente, a contrario, un enrichissement extrêmement fort. Il apparaît donc légitime de douter de la véracité de ces résultats de ChIP. Plusieurs modifications du protocole de ChIP ont été introduites pour réduire les interactions aspécifiques et d'autres contrôles négatifs comme GAPDH et 36B4 ont été étudiés.



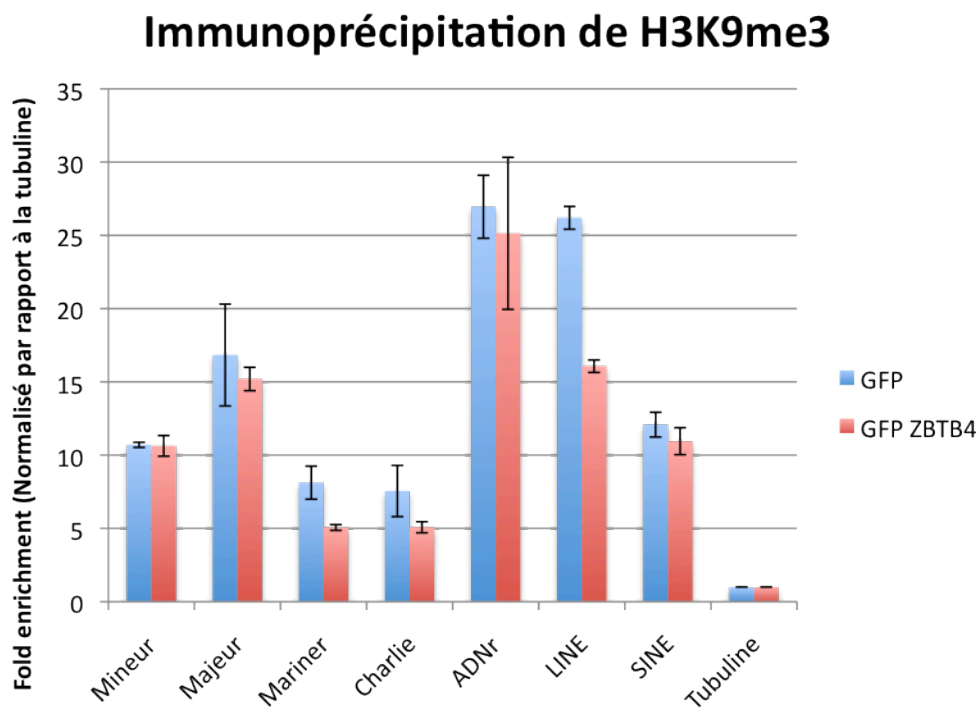
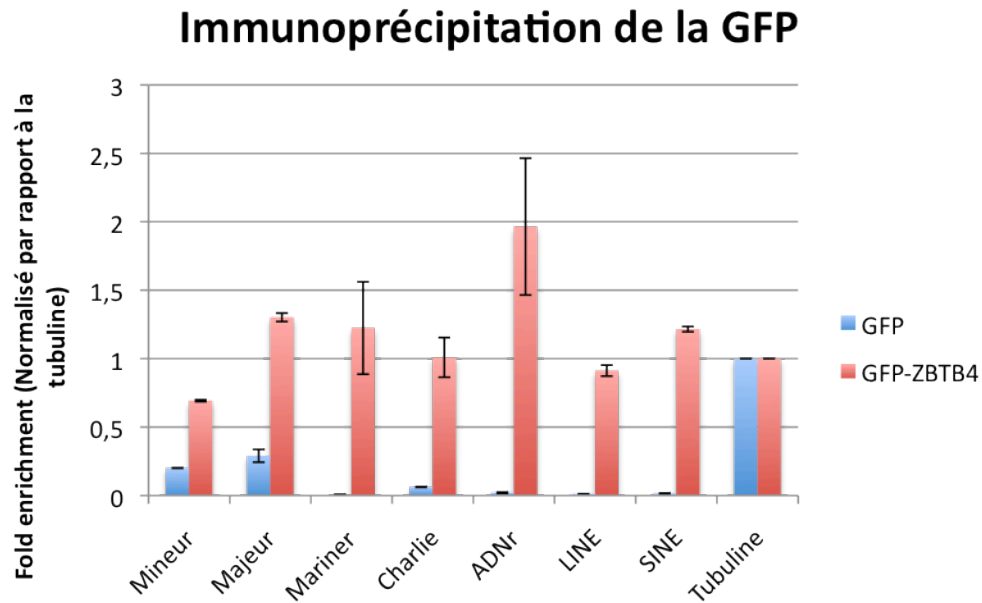


**Figure 35. Analyse d'expression des éléments répétés par PCR quantitative.**

L'expression de différents éléments répétés a été analysée par PCR quantitative dans le foie et les reins de souris sauvages et mutantes pour ZBTB4. L'expression des satellites mineurs est augmentée de 2,5 fois dans le foie des souris mutantes pour ZBTB4. Les barres d'erreurs correspondent aux erreurs standard des triplicats techniques. Les données d'expression ont été normalisées par rapport à GAPDH.



**Figure 36. Immunoprécipitation des satellites mineurs et majeurs avec l'anticorps ZBTB4 #120**  
 ZBTB4 ne semble pas enrichie au niveau des satellites mineurs et majeurs en comparaison avec les contrôles négatifs que sont les billes et les immunoglobulines, tandis que l'on détecte bien un enrichissement de la marque H3K9me3, contrôle positif, au niveau de ces éléments répétés.



**Figure 37. Immunoprécipitation des éléments répétés après surexpression de ZBTB4.**

ZBTB4 est légèrement enrichie au niveau des satellites majeurs et de l'ADNr. La marque H3K9me3, comme contrôle positif, est enrichie sur ces différents éléments répétés, notamment sur les satellites mineurs et majeurs, l'ADNr et les éléments LINE et SINE. Les enrichissements sont mesurés par qPCR. Les barres d'erreurs correspondant à l'erreur standard des triplicats techniques. Les données ont été normalisées par rapport à la tubuline.

## 2.3 Discussion

Les résultats préliminaires d'analyses d'expression semblent indiquer une réexpression partielle des satellites mineurs, caractéristiques des régions centromériques chez les souris *Zbtb4*<sup>-/-</sup>. Cependant, cette augmentation d'expression chez les mutants est faible, comme l'atteste les résultats de PCR semi-quantitative qui ne permettent pas de détecter un changement drastique d'expression. Par ailleurs, la réexpression semble se produire uniquement dans le foie et l'absence de *Zbtb4* ne perturbe pas les autres organes testés. De la même manière, l'expression des éléments Charlie, LINE et SINE est restreinte aux reins. Une explication possible à ces observations est que d'autres MBPs, en particulier les protéines Kaiso et *Zbtb38*, seraient capables de compenser la perte de *Zbtb4* de façon spécifique à certains tissus. Cette hypothèse permettrait également d'expliquer la variabilité entre les souris. En effet, lors de la génération d'une cohorte pour un phénotypage à l'Institut Clinique de la Souris, les résultats du génotypage de 482 mâles montre qu'il y a uniquement 20% de souris *Zbtb4*<sup>-/-</sup> viables contre 25% attendus. La létalité embryonnaire des souris délétées pour *Zbtb4* serait alors compensée par la présence d'autres MBPs ayant une fonction redondante en étant recrutées sur des cibles communes. Cette hypothèse sera testée par Laetitia Kasprzyk dans le cadre de son projet de caractérisation de la souris mutante.

Les expériences de ChIP ne nous permettent pas de valider ces différents éléments répétés comme étant des cibles directes pour *Zbtb4* à cause d'un enrichissement élevé sur des cibles non spécifiques comme la tubuline. Bien qu'il fut intéressant de transfecter la GFP-ZBTB4 afin de déterminer l'efficacité de transfection, il sera nécessaire par la suite de reproduire ces expériences en utilisant un plasmide ZBTB4 étiqueté par un FLAG, plus efficace que la GFP. Par ailleurs, la fabrication d'un anticorps murin dirigé contre *Zbtb4* est en cours et nous permettra alors d'étudier la localisation de la protéine endogène.

De manière générale, nous chercherons à montrer que *Zbtb4* contrôle l'expression des satellites mineurs afin de réguler cette structure indispensable à la ségrégation des chromosomes au cours de la mitose et de

la méiose. C'est en effet au niveau du centromère que se forme le kinétochore, structure protéique responsable de l'attachement des microtubules, permettant le mouvement des chromosomes dans les cellules filles (Vagnarelli et al. 2008). Sachant cela, nous étudieront la présence éventuelle de défauts de ségrégation au cours de la mitose dans les souris *Zbtb4*<sup>-/-</sup>. La présence de réarrangements chromosomiques pourra être déterminée par l'observation de plaques métaphasiques et nous analyseront si la réexpression des satellites mineurs, dans les cellules *Zbtb4*<sup>-/-</sup>, induit une instabilité des centromères par l'observation du nombre et de la taille des chromocentres. Par ailleurs, l'instabilité centromériques pourrait causer un réarrangement du nombre de copie des satellites mineurs qui pourra être quantifié par qPCR à partir d'ADN génomique de souris sauvages et mutantes pour *Zbtb4*.

D'un point de vue épigénétique, il sera également intéressant de savoir si la perte de *Zbtb4* provoque un changement d'état chromatinien au niveau des satellites mineurs. Pour cela, des expériences d'immunofluorescence et de ChIP permettront de déterminer le statut de méthylation de la lysine 9 de l'histone H3 normalement triméthylé au niveau des régions centromériques (Lehnertz et al. 2003 ; Martens et al. 2005). Il sera également intéressant de vérifier si les souris mutantes pour *Zbtb4* affichent une hypométhylation spécifique des satellites mineurs par des expériences séquençages au bisulfite ou par digestion de l'ADN avec des enzymes sensibles à la méthylation.

## 2.4 Matériels et méthodes

### Anticorps et plasmide

Le plasmide ZBTB4 fusionné à la GFP a été généré précédemment par notre laboratoire (Filion et al. 2006). La protéine entière a été clonée en C-terminal de la GFP sous un promoteur CMV dans le plasmide pEGFP-C2. La liste des anticorps utilisés au cours de cette étude est répertoriée table 18.

**Table 18. Anticorps**

Anticorps	Source
ZBTB4 #120	À façon
ZBTB4 #98	À façon
ZBTB4 SC	Santa Cruz (SC-161265)
H3K9me3	Abcam (ab8898)

#### Extraction d'ARN à partir de tissus

Cent mg de tissus sont homogénéisés dans 1 ml de Trizol Reagent (*Invitrogen*) à l'aide d'un pilon et d'un mortier après congélation des tissus à l'azote liquide. Les échantillons homogénéisés sont incubés 5 min à température ambiante pour permettre la dissociation des complexes nucléoprotéiques. Deux cents µl de chloroforme sont ajoutés et les tubes sont incubés 3 min à température ambiante puis centrifugés à 12 000 x g pendant 15 min à 4°C. Après centrifugation, le mélange est divisé en deux phases distinctes séparées par une interphase. L'ARN se trouve dans la phase aqueuse supérieure qui est alors récupérée dans un nouveau tube de 1,5 ml. Cinq cents µl sont ajoutés à la phase aqueuse et les échantillons sont incubés à température ambiante pendant 10 min puis centrifugés à 4°C pendant 10 min à 12 000 x g. L'ARN forme un précipité et le surnageant est retiré. L'ARN est ensuite lavé par l'ajout de 1 ml d'alcool 75 % et centrifugé à 7 500 x g pendant 5 min à 4°C. Après centrifugation, le surnageant est retiré et le précipité séché puis resuspendu dans 30 µl d'eau. La concentration d'ARN est alors déterminée au Nanodrop (*Thermo Scientific*).

#### Digestion à la DNase

Afin d'éliminer une contamination par l'ADN un traitement à la DNase est nécessaire. Deux µg d'ARN sont incubés avec 1 µl de Turbo DNase (1 unité ; *Ambion*) et 1 µl de tampon 10X Turbo DNase et complétés à 10 µl avec de l'eau. La réaction est incubée à 37°C pendant 30 min. La Turbo DNase est alors inactivée par l'ajout de 2 µl de réactif d'inactivation (*Ambion*), incubée pendant 5 min à température ambiante. La réaction est

centrifugée à 10 000 x g pendant 2 min permettant la précipitation du réactif d'inactivation. Après centrifugation, le surnageant contenant l'ARN est récupéré dans un nouveau tube.

#### Réaction de rétrotranscription

Les 10 µl de la réaction de digestion à la DNase, soit environ 2 µg d'ARN, sont alors rétrotranscrits en ADNc grâce au kit SuperScript VILO cDNA synthesis (*Invitrogen*). Pour cela, les 10 µl sont incubés avec 4 µl de mix 5X VILO Reaction, 2 µl de mix 10X SuperScript Enzyme et 4 µl d'eau. La réaction est incubée à 25°C pendant 10 min puis à 42°C pendant 60 min et enfin à 85°C pendant 5 min.

#### Transfection de cellules de mammifères

Les cellules de souris MEFs (Mouse Embryonic Fibroblast) sont cultivées en DMEM (*Gibco*) 10% SVF (sérum de veau fœtal) sans antibiotique à 37°C. La veille des transfections, les cellules sont trypsinées et ensemencées dans des boîtes de Pétri de 10 cm pour qu'elles soient confluentes à 80-85% au moment des transfections. La transfection est réalisée par la Lipofectamine 2000 (*Invitrogen*). Chaque puit est transfecté par 24 µg de plasmide et 60 µl de Lipofectamine 2000 dilués dans 3 ml de milieu Opti-MEM (*Gibco*). 4 à 6 heures après la transfection le milieu est remplacé par du DMEM 10% SVF sans antibiotique. Les cellules sont récoltées après 21 heures de transfection.

#### Extraction nucléaire

Après 21 heures de transfection, les cellules sont trypsinées et transférées dans un tube de 50 ml. Les cellules sont précipitées par centrifugation à 500 x g pendant 3 min et resuspendues, à une concentration de 10<sup>7</sup> cellules par tube, dans 10 ml de MEM (*Gibco*) avec 1% de formaldéhyde pendant 10 min sous agitation. Après 10 min d'incubation, la réaction de fixation est arrêtée par l'ajout de 500 µl de glycine 2,5 M et incubée pendant 2 min à température ambiante puis les cellules sont centrifugées à 500 x g pendant 3 min. Le milieu est alors retiré et les cellules sont lavées une fois avec 10 ml de PBS puis resuspendues dans 100 µl de tampon B de sonication (*Diagenode*).

### Sonication

Cent µl de tampon B (*Diagenode*) contenant  $10^6$  cellules sont soniquées en présence de 5 µl d'inhibiteur de protéases (*Diagenode*). Vingt cycles de 30s ON et 30s OFF sont réalisées pour un temps total de 20 min. La chromatine soniquée est alors diluée dans 900 µl de tampon A (*Diagenode*).

### Immunoprécipitation de la chromatine

Les expériences de ChIP sont réalisées à l'aide du robot SX-8G IP Star (*Diagenode*). Un million de cellules sont utilisées pour effectuer une immunoprécipitation. Un µg d'anticorps est dilué dans 100 µl de tampon A (*Diagenode*) et 10 µl de billes magnétiques (*Diagenode*) ainsi que les différents tampons de lavage et d'élution sont préparées dans des barrettes de 12 tubes de 200 µl. Le temps d'immobilisation des billes avec les anticorps est ajusté à 5 heures tandis que la réaction d'immunoprécipitation a une durée de 10 heures. En fin de réaction, 1 µl de protéinase K (*Diagenode*) est ajouté à l'input et aux éluats puis incubé à nouveau dans l'automate qui procède à la réaction de decrosslink à 65°C pendant 4 heures. Les échantillons sont alors directement analysés par qPCR.

### Séquences

Les séquences des amorces utilisées pour les PCR semi-quantitatives et quantitatives sont listées table 19.



**Table 19. Amorces utilisées au cours des différentes PCR.**

Oligo	Gène	Séquence	Commentaire
MJ136	Satellites majeurs	GACGACTTGAAAAATGACGAAATC	Amorce 5'
MJ137	Satellites majeurs	CATATTCCAGGTCCTTCAGTGTGC	Amorce 3'
MJ138	Satellites mineurs	CATGGAAAATGATAAAAACC	Amorce 5'
MJ139	Satellites mineurs	CATCTAATATGTTCTACAGTGTGG	Amorce 3'
MJ140	Mariner	AGGCAGCAGAGCACAAATG	Amorce 5'
MJ141	Mariner	TTGCTGTTAAGGGAATTGTGG	Amorce 3'
MJ142	Charlie	TTGAGAATCGGATGGGAGAC	Amorce 5'
MJ143	Charlie	AAGAACTGTCTTATTCAGGC	Amorce 3'
MJ146	IAP	AGCAGGTGAAGCCACTG	Amorce 5'
MJ147	IAP	CTTGCCACACTTAGAGC	Amorce 3'
MJ150	LINE	TTTGGGACACAATGAAAGCA	Amorce 5'
MJ151	LINE	CTGCCGTCTACTCCTCTTGG	Amorce 3'
MJ154	SINE B2	GAGATGGCTCAGTGGTTAAG	Amorce 5'
MJ155	SINE B2	CTGTCTTCAGACACTCCAG	Amorce 3'
MJ156	ADNr	CCTGTGAATTCTCTGAATC	Amorce 5'
MJ157	ADNr	CCTAACTGCTGACAGGGTG	Amorce 3'
MJ 158	Tubuline	GACAGAGGCAAACCTGAGCACC	Amorce 5'
MJ159	Tubuline	CAACGTCAAGACGGCCGTGTG	Amorce 3'





## Partie IV : conclusion générale

L'épigénétique est un composant essentiel du fonctionnement des génomes eucaryotes. Les divers phénomènes épigénétiques modifient l'état chromatinien et participent à la plasticité du génome, mais aussi au maintien de son identité fonctionnelle à travers les générations cellulaires. Parmi ces processus, la méthylation de l'ADN joue un rôle fondamental dans la régulation de l'expression des gènes.

La méthylation exerce ces effets par l'intermédiaire de protéines dédiées, qui reconnaissent spécifiquement l'ADN méthylé et contrôlent la transcription en modulant la chromatine. Trois familles de protéines liant l'ADN méthylé sont connues chez les mammifères, et elles totalisent entre elles neuf membres. De nombreux arguments suggèrent que cette liste est encore incomplète, et que des protéines humaines liant l'ADN méthylé restent à découvrir.

Par deux approches distinctes, nous avons essayé d'identifier de nouvelles protéines humaines se liant à l'ADN méthylé. En utilisant des données bibliographiques récentes, nous avons établi une liste de six protéines candidates, dont nous avons testé la capacité à lier l'ADN méthylé par deux approches complémentaires. La première vise à déterminer la localisation intracellulaire des protéines *in vivo*, pour déterminer si elles colocalisent avec les régions du génome connues pour être méthylées. La deuxième consiste à étudier l'interaction entre protéines et ADN méthylé *in vitro* par des expériences de retard sur gel.

Nous avons achevé l'ensemble de ces expériences. Les contrôles positifs et négatifs inclus permettent de tirer une conclusion sans équivoque : les protéines candidates étudiées ne lient pas l'ADN méthylé dans nos conditions expérimentales.

Nous avons alors mis en place une nouvelle stratégie reposant sur le "phage-display" : une banque de phages, exprimant à leur surface des protéines de fusion codées par des ADNc humains, est mise en contact avec des billes magnétiques recouvertes d'ADN méthylé. Elle est rapide et

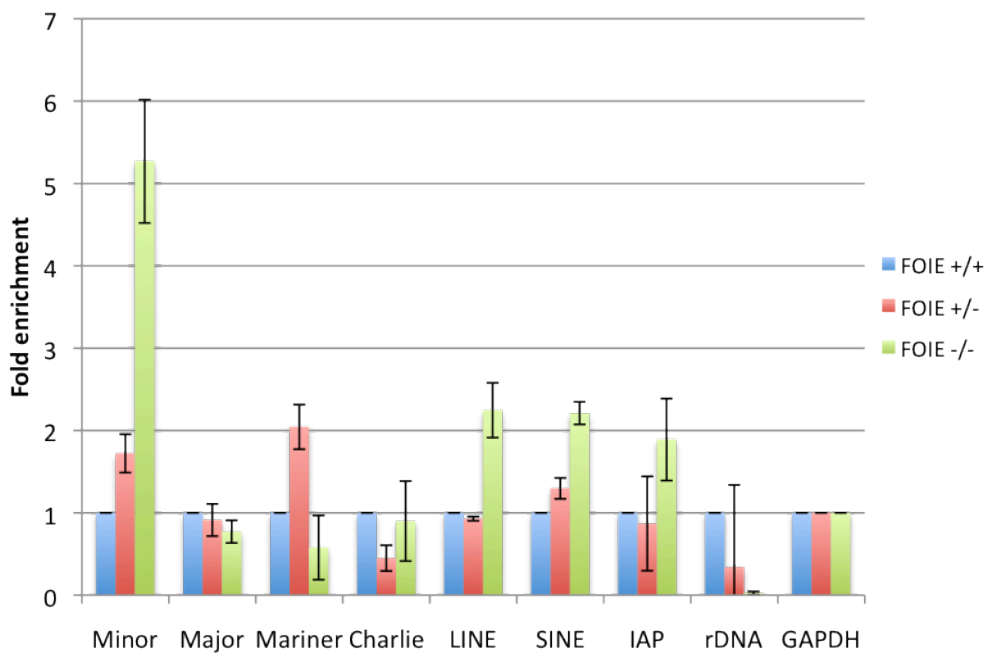
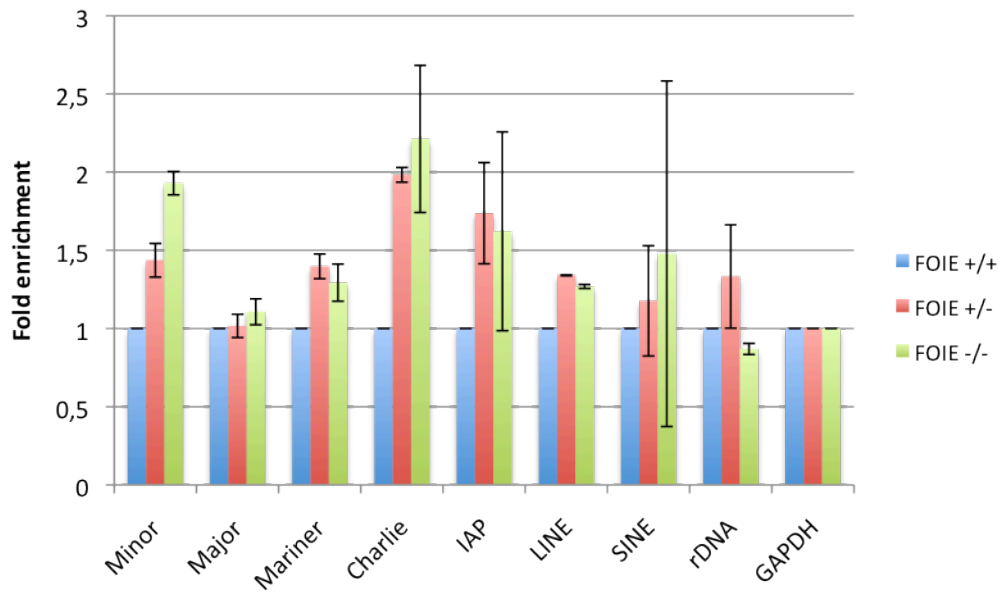
flexible : les séquences méthylées peuvent facilement être variées, de même que les conditions physico-chimiques d'incubation ADN-protéine.

Deux candidats, CHD3 et HMGB1, sélectionnés pour lier l'ADN méthylé se révèlent intéressants de par leur fonction, respectivement, dans le maintien des centrosomes et le remodelage de l'ADN. Des études supplémentaires sont désormais nécessaires afin de valider ces candidats.





## Annexe 1



### Analyse d'expression des différents éléments répétés dans le foie par qPCR.

L'expression de différents éléments répétés a été analysée par PCR quantitative dans le foie de souris sauvages et mutantes pour ZBTB4. L'expression des satellites mineurs est augmentée, respectivement, de 2 et 5 fois dans le foie des souris mutantes pour ZBTB4. Les barres d'erreurs correspondent aux erreurs standard des triplicats techniques. Les données d'expression ont été normalisées par rapport à GAPDH.









## Annexe 2

Mammalian methyl-binding proteins:  
What might they do?

Michael Joulie, Benoit Miotto et Pierre-Antoine Defossez.

BioEssays, Décembre 2010



### **Annexe 3**

#### Définition de l'épigénétique

Michael Joulie et Pierre-Antoine Defossez

La Science au présent, 2010



# L'épigénétique.

Michael Joulie, Pierre-Antoine Defossez  
CNRS UMR7216 "Epigénétique et Destin Cellulaire"  
Université Paris Diderot, Bâtiment Lamarck, case 7042.  
35 rue Hélène Brion, 75205 Paris CEDEX 13.  
Email : Pierre-Antoine.Defossez@univ-paris-diderot.fr

Le 20<sup>ème</sup> siècle a été marqué par la génétique. Ses avancées techniques et conceptuelles sont extraordinaires, mais ne suffisent pas à expliquer toute la biologie de l'ADN. Des découvertes récentes montrent qu'un ensemble de phénomènes, dits "épigénétiques" agissent sur notre génome. Ils assurent son bon fonctionnement dans l'espace et le temps, mais peuvent également connaître des dysfonctionnements et causer des maladies.

## Avancées et limites de la génétique

Rappelons ici que l'ADN porte l'identité génétique, qu'il peut subir des mutations, et que l'information qu'il contient est structurée en unités distinctes, les gènes. Ceux-ci sont véhiculés par les chromosomes. Par la transcription, les gènes donnent naissance à des molécules d'ARN. Celles-ci peuvent avoir un rôle intrinsèque. Mais elle peuvent aussi être converties en protéines, par le phénomène de traduction. Les gènes ne sont pas toujours actifs, mais peuvent, selon les conditions, être transcrits ou non. Ceci dépend de l'action de protéines régulatrices, les facteurs de transcription, qui peuvent stimuler ou ralentir leur expression. De même, la traduction d'un ARN en protéine n'est pas systématique, mais est régulable. Le génome humain a été séquencé et nous connaissons maintenant le nombre et la position de presque tous nos gènes.

On pourrait croire que ces connaissances suffisent à comprendre le fonctionnement de notre génome, mais ce n'est pas le cas. Différents exemples (voir illustration 1, encadrés, et texte plus bas), montrent que la séquence de l'ADN n'est pas le seul élément déterminant son activité : d'autres facteurs agissent. Certains sont "épigénétiques" c'est à dire qu'ils affectent l'activité du génome de façon héritable, mais ne modifient pas sa séquence. Donnons une analogie : le génome peut porter des marques épigénétiques qui modifient sa lecture sans changer sa séquence, de la même façon qu'un texte peut être souligné ou surligné sans que les lettres qui le constituent ne soient changées.



# Les phénomènes épigénétiques

## Méthylation de l'ADN

La méthylation est l'étiquetage de l'ADN par un groupement chimique, le méthyle ou CH<sub>3</sub>. Différentes enzymes posent cette marque, et la renouvellent au cours des divisions cellulaires. En général, la méthylation de l'ADN réprime l'expression des gènes, par deux phénomènes parallèles. D'abord, la méthylation modifie la surface de l'ADN, et peut empêcher sa reconnaissance par certaines protéines dont le rôle est d'activer l'expression des gènes. De plus, l'ADN méthylé est reconnu par certaines protéines spécialisées qui s'y fixent et créent localement un environnement s'opposant à la transcription.

## Dynamique des histones

Chaque cellule humaine contient deux mètres d'ADN! Pour loger dans un noyau de quelques microns, l'ADN est rendu plus compact par son enroulement autour de protéines, appelées histones.

Les histones présentent trois niveaux de complexité. 1) il existe différents types d'histones. 2) les histones ne sont pas figées : elles peuvent être déplacées le long de l'ADN, ou ôtées. 3) les histones peuvent être marquées par diverses modifications chimiques comme la phosphorylation ou l'acétylation.

La nature précise des histones, leur position, et le patron de leurs modifications influent directement sur l'activité des gènes localisés à leur proximité. De plus, différents mécanismes assurent la maintenance de ces différents paramètres au cours du temps et des divisions cellulaires.

## ARNs non codants

Les ARNs peuvent être "codants", c'est à dire servir d'intermédiaires de l'information entre ADN et protéines. Mais ils peuvent avoir une activité intrinsèque, comme les ARNs composant les ribosomes. Ils sont alors dits "non-codants". Un ensemble de travaux, récompensé par le Prix Nobel en 2006, a montré que les ARNs non-codants sont très nombreux, et qu'ils agissent sur l'expression des gènes à différents niveaux. Ils peuvent agir sur des ARNs codants, pour les déstabiliser ou inhiber leur traduction. Mais ils peuvent aussi exercer une action directement sur l'ADN.

Différents ARNs non-codants s'apparient avec les régions du génome partageant leur séquence, et provoquer sur ces cibles la formation d'un environnement réprimant l'expression des gènes. Ce phénomène passe par la modification des histones et la méthylation de l'ADN, évoqués plus haut. De façon plus générale, les processus épigénétiques sont fortement reliés, puisque la méthylation de l'ADN peut causer la modification des histones, et vice-versa.

Certains ARNs non-codants agissant sur le génome sont petits (une vingtaine de nucléotides), d'autre beaucoup plus grands (des milliers de nucléotides). C'est le

cas de l'ARN Xist, dont les nombreuses copies recouvrent et inactivent un chromosome entier chez les mammifères femelles (illustration 1).

## **Conséquences pratiques de l'épigénétique**

### **Epigénétique et cancer**

Le cancer peut avoir une base génétique, si un gène est touché par une mutation, une délétion, ou une translocation chromosomique. Mais on sait désormais que les anomalies épigénétiques sont au moins aussi importantes.

Les mieux connues touchent la méthylation de l'ADN. Pour des raisons encore obscures, certains gènes anti-tumoraux peuvent se voir méthylés alors qu'ils ne devraient pas l'être. Ceci cause leur inactivation, et rend plus facile la transformation tumorale de la cellule affectée. Ce phénomène est très fréquemment observé, et de nouvelles thérapies ont été développées pour s'y opposer en diminuant la méthylation de l'ADN. Elles donnent des résultats très encourageants, en particulier contre des leucémies.

### **Empreinte parentale, maladies héréditaires**

Certaines maladies sont héréditaires, mais ont une transmission complexe. On connaît ainsi une région précise du chromosome 15, dont la délétion provoque des maladies mentales. Si une femme a cette délétion, ses enfants héritant du chromosome anormal auront un syndrome d'Angelman, caractérisé par un retard mental, des mouvements saccadés ainsi qu'un comportement extrêmement joyeux. Si un homme possède exactement la même délétion, et qu'il la transmet à ses enfants, ceux-ci développeront le syndrome de Prader Willi, avec notamment un comportement compulsif vis-à-vis de la nourriture donnant une obésité morbide.

Une même anomalie génétique donne donc deux maladies distinctes, selon le parent affecté. La raison de cette asymétrie est le phénomène d'empreinte parentale : la région touchée par la délétion contient des gènes qui sont exprimés différemment sur le chromosome paternel ou maternel. Leur manque a donc des conséquences phénotypiques différentes. L'empreinte parentale est régie par la méthylation de l'ADN, les modifications des histones, et les ARNs non codants.

### **Cellules souches et reprogrammation nucléaire**

La médecine porte beaucoup d'espoirs sur les cellules souches. Ce sont des cellules pluripotentes que l'on peut cultiver in vitro, puis réimplanter dans le corps pour remplacer des cellules déficientes. Leurs applications potentielles sont importantes, comme renouveler les neurones dans les maladies d'Alzheimer ou de Parkinson.

Il est difficile d'isoler des cellules souches à partir de donneurs car elles sont peu abondantes et difficiles d'accès. Il semble plus aisé de les re-crée in vitro. Pour cela on prend des cellules différenciées mais facilement accessibles, par exemple des cellules de la peau. Il faut alors les "re-programmer" pour qu'elles retrouvent le caractère pluripotent qu'elles ont perdu. Une des difficultés principales est de

reprogrammer correctement le programme épigénétique des cellules. De nombreux laboratoires visent à caractériser précisément l'état épigénétique des cellules souches, celui des cellules différenciées, et à les reprogrammer plus efficacement.

### **Épigénétique et environnement**

Les radiations nucléaires, les rayons ultraviolets, mais aussi certaines molécules présentes dans notre environnement sont mutagènes : ils peuvent altérer l'ADN et changer sa séquence. Différents tests existent pour identifier ces produits et nous en protéger.

Des études de laboratoire ont montré que certains pesticides ont des effets épigénétiques chez le rat : ils altèrent la méthylation de l'ADN et diminuent la fertilité des mâles, génération après génération, sans pour autant causer de mutations. Très certainement, d'autres molécules peuvent perturber le paysage épigénétique de nos cellules, mais nous n'en connaissons que peu, et ne disposons pas de tests systématiques pour les identifier. Il faudra dans le futur dresser la liste de ces "épimutagènes".

L'environnement apporte aussi des molécules nécessaires à notre machinerie épigénétique. Ainsi le folate, un nutriment apparenté à la vitamine B9 fourni par notre alimentation, est nécessaire pour méthyler l'ADN. Un manque de folate conduit donc à une méthylation réduite de l'ADN. Il en existe des exemples historiques chez l'homme : au cours de la 2ème guerre mondiale, la famine a atteint les Pays-Bas; les femmes alors enceintes donnèrent naissance à des enfants de petite taille. Ces enfants, devenus adultes, eurent eux-mêmes des descendants. Contre toute attente, ces descendants étaient également petits alors qu'ils n'avaient pas subi de restrictions. Des recherches ont montré que la famine avait provoqué un manque de folate chez les femmes enceintes, causant des anomalies épigénétiques chez leurs enfants, qui se perpétuent chez leurs petits-enfants.

### **Conclusion et perspectives**

Lamarck proposait que l'environnement influence directement nos caractéristiques, et que les changements produits soient fixés dans notre descendance. Après un siècle très darwinien, il s'avère que l'environnement peut effectivement avoir un effet direct sur notre génome, et que son action peut se perpétuer à travers les générations, même après la disparition du stimulus initial. Une petite dose de Lamarckisme est réintroduite dans la biologie et l'évolution.

On comprend aujourd'hui que la séquence d'ADN est un jeu d'instructions nécessaire, mais non suffisant, pour le développement d'un organisme complexe. Des informations complémentaires, portées par l'ADN sous forme de méthylation, mais aussi par les protéines et les ARNs associés à l'ADN, sont aussi requises pour le bon fonctionnement de notre génome, et donc la fonction de nos cellules.

En dehors de la progression des connaissances fondamentales, on peut attendre des conséquences pratiques de ces avancées, dans le domaine du diagnostic

médical, du développement de médicaments, et de la production de cellules souches.

Références :

Histoire de la biologie moléculaire, Michel Morange. La Découverte, 2003.

Jean-Paul Renard, le clonage, Encyclopedia Universalis, La Science au présent 2003.

Médecine/Sciences, numéro spécial "Epigénétique", Août-Septembre 2008.

## **Encadré 1 : Empreinte parentale**

Chaque cellule humaine contient deux copies de chaque gène. L'une est portée par le chromosome apporté par le père, et l'autre par le chromosome d'origine maternelle. Dans la plupart des cas, ces deux copies sont exprimées de manière équivalente. Une centaine de gènes au moins échappent à cette règle et sont exprimés uniquement par l'un ou l'autre des chromosomes, et toujours le même (soit le chromosome d'origine paternelle, soit celui d'origine maternelle). On dit de ces gènes qu'ils sont "soumis à l'empreinte parentale". Beaucoup d'entre eux agissent lors du développement embryonnaire, où ils régulent la croissance du placenta ou de l'embryon. Chez les animaux, l'empreinte est une "invention" relativement récente au cours de l'évolution : elle n'existe que chez les marsupiaux et les mammifères placentaires. Les raisons de sa sélection évolutive sont débattues, mais deux hypothèses non-exclusives prévalent. Premièrement, l'existence de l'empreinte empêche la parthénogenèse (développement d'un embryon à partir d'un oeuf non fécondé). Deuxièmement, l'empreinte reflète les intérêts évolutifs divergents des mammifères mâles et femelles. En effet, il est positif pour le père que l'embryon ayant son génome maximise sa propre croissance. Pour la mère, qui peut porter plusieurs embryons ayant des pères différents, il est au contraire bénéfique de modérer cette croissance et de répartir les ressources nutritives entre les embryons. L'empreinte est un phénomène épigénétique, qui repose sur la méthylation de l'ADN, les modifications des histones, et l'utilisation d'ARNs non-codants.

## **Encadré 2 : Les souris Agouti**

Chez la souris, le gène *Agouti* est responsable de la couleur du pelage. Lorsque celui-ci est actif, les souris sont de couleur jaune. En revanche, s'il n'est pas exprimé, elles sont brunes et tachetées. Il s'avère que l'expression du gène *Agouti* est régulée par la méthylation de l'ADN. La méthylation du gène inhibe son expression conduisant à un caractère tacheté du pelage. Inversement, l'absence de méthylation induit l'expression du gène donnant alors la couleur jaune au pelage. On sait désormais que le régime alimentaire de la mère influence la méthylation du génome. Si l'alimentation de la souris gestante est riche en précurseur de méthyle, comme le folate, alors les souriceaux ont tendance à avoir un pelage brun et tacheté. En revanche, une nourriture pauvre en acide folique donne un pelage clair. Le faible taux de méthylation des souris jaunes se retrouve non seulement

sur le gène *Agouti* mais aussi sur le reste du génome. Cette hypométhylation globale du génome a diverses conséquences chez ces souris comme une plus grande susceptibilité face au diabète, à l'obésité ou au cancer.

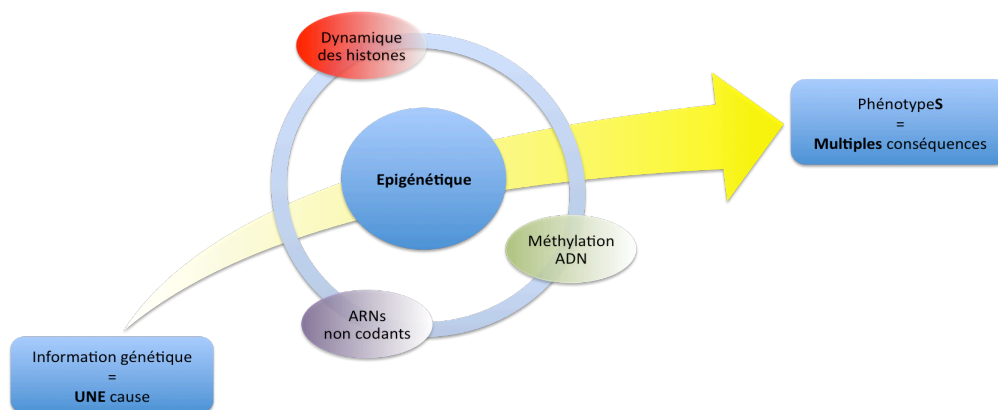
Les chattes calico ont un pelage bigarré orange et noir. Toutes leurs cellules contiennent les mêmes chromosomes, mais certaines produisent un pigment orange, d'autres un pigment noir. Comment l'expliquer?

Il faut savoir que le gène responsable de la couleur du pelage se trouve sur le chromosome X. Ce gène a deux versions (ou allèles). Un allèle donne la couleur noire tandis que l'autre donne la couleur orange.

L'autre élément expliquant le phénomène est le suivant. Les mammifères mâles et femelles diffèrent par leurs chromosomes sexuels. Les mâles ont un X et un Y; les femelles deux X. Le chromosome X contient de nombreux gènes, et on pourrait imaginer que ces gènes soient exprimés deux fois plus chez les femelles que chez les mâles. Ce n'est pas le cas, car il existe un mécanisme de compensation, appelé inactivation du chromosome X. Dans toutes les cellules femelles, un des deux chromosomes X est rendu totalement silencieux, par un phénomène épigénétique impliquant ARNs non codants, modifications des histones, et méthylation de l'ADN. Le choix initial du chromosome inactivé est aléatoire. Une fois inactivé, un chromosome reste inactif dans toute la descendance de la cellule initiale.

On peut donc expliquer l'apparence de la chatte calico comme suit. Dans l'embryon, chaque cellule contribuant au pelage "choisit" d'inactiver le gène donnant un pigment noir, ou le gène donnant un pigment orange. Ce "choix" est maintenu lors de la division des cellules, donnant des plaques oranges ou noires.

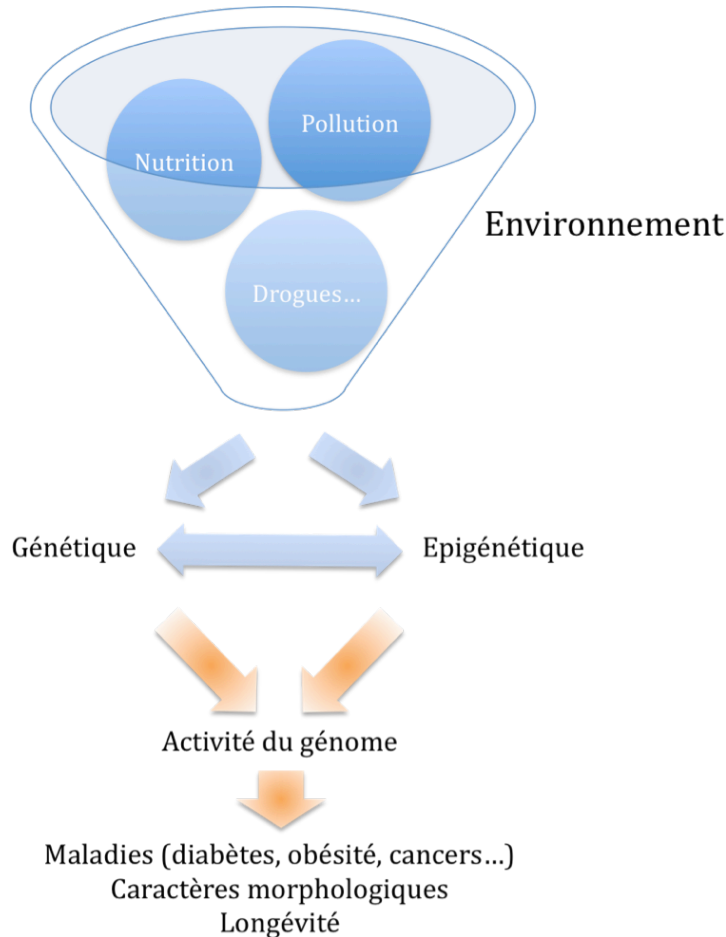
## Illustration 2 :



Les mécanismes épigénétiques concernent les histones, les ARNs non codants, et la méthylation de l'ADN. Ces phénomènes modulent la lecture de notre information génétique. Il s'ensuit que nos différentes cellules ont bien le même patrimoine génétique, mais pas nécessairement le même phénotype, et que des individus ayant le même génome peuvent avoir des caractéristiques bien différentes. L'épigénétique apporte un degré supplémentaire de régulation du génome, et son existence permet la

génération d'organismes multicellulaires complexes et réactifs à leur environnement.

### Illustration 3 :



L'environnement influence notre génome de plusieurs façons. Au niveau génétique, on connaît un grand nombre de facteurs mutagènes capable d'induire des mutations provoquant des cancers (U.V, tabac...). Avec l'émergence de l'épigénétique, les chercheurs s'aperçoivent que l'environnement peut non seulement agir sur un plan génétique mais également sur un plan épigénétique. Les facteurs environnementaux vont perturber l'activité du génome via les mécanismes épigénétiques induisant des caractères anormaux.



## Bibliographie

- Achour, M., Fuhrmann, G., Alhosin, M., Rondé, P., Chataigneau, T., Mousli, M., et al. (2009). UHRF1 recruits the histone acetyltransferase Tip60 and controls its expression and activity. *Biochemical and Biophysical Research Communications*, 390 (3), 523-8.
- Achour, M., Jacq, X., Rondé, P., Alhosin, M., Charlot, C., Chataigneau, T., et al. (2008). The interaction of the SRA domain of ICBP90 with a novel domain of DNMT1 is involved in the regulation of VEGF gene expression. *Oncogene*, 27 (15), 2187-97.
- Agarwal, N., Hardt, T., Brero, A., Nowak, D., Rothbauer, U., Becker, A., et al. (2007). MeCP2 interacts with HP1 and modulates its heterochromatin association during myogenic differentiation. *Nucleic Acids Research*, 35 (16), 5402-5408.
- Allan, A., Liang, X., Luo, Y., Pak, C., Li, X., Szulwach, K., et al. (2008). The loss of methyl-CpG binding protein 1 leads to autism-like behavioral deficits. *Human molecular genetics*, 17 (13), 2047-57.
- Amir, R., Van den Veyver, I., Wan, M., Tran, C., Francke, U., & Zoghbi, H. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics*, 23 (2), 185-8.
- Angrisano, T., Lembo, F., Pero, R., Natale, F., Fusco, A., Avvedimento, V., et al. (2006). TACC3 mediates the association of MBD2 with histone acetyltransferases and relieves transcriptional repression of methylated promoters. *Nucleic Acids Research*, 34 (1), 364-72.
- Aoki, A., Suetake, I., Miyagawa, J., Fujio, T., Chijiwa, T., Sasaki, H., et al. (2001). Enzymatic properties of de novo-type mouse DNA (cytosine-5) methyltransferases. *Nucleic acids research*, 29 (17), 3506-12.
- Aravin, A., & Bourc'his, D. (2008). Small RNA guides for de novo DNA methylation in mammalian germ cells. *Genes & Development*, 22 (8), 970-5.
- Aravin, A., Sachidanandam, R., Girard, A., Fejes-Toth, K., & Hannon, G. (2007). Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science (New York, NY)*, 316 (5825), 744-7.
- Auriol, E., Billard, L.-M., Magdinier, F., & Dante, R. (2005). Specific binding of the methyl binding domain protein 2 at the BRCA1-NBR2 locus. *Nucleic Acids Research*, 33 (13), 4243-54.
- Ball, M., Li, J., Gao, Y., Lee, J.-H., LeProust, E., Park, I.-H., et al. (2009). Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nature biotechnology*, 27 (4), 361-8.
- Bannister, A., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell research*, 21 (3), 381-95.
- Bannister, A., Zegerman, P., Partridge, J., Miska, E., Thomas, J., Allshire, R., et al. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*, 410 (6824), 120-124.
- Barbas, C., Kang, A., Lerner, R., & Benkovic, S. (1991). Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proceedings of the National Academy of Sciences of the United States of America*, 88 (18), 7978-82.
- Bartke, T., Vermeulen, M., Xhemalce, B., Robson, S., Mann, M., & Kouzarides, T. (2010). Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell*, 143 (3), 470-84.
- Beisel, C., & Paro, R. (2011). Silencing chromatin: comparing modes and mechanisms. *Nature reviews Genetics*, 12 (2), 123-35.
- Bellacosa, A., Cicchillitti, L., Schepis, F., Riccio, A., Yeung, A., Matsumoto, Y., et al. (1999). MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. *Proceedings of the National Academy of Sciences of the United States of America*, 96 (7), 3969-74.



Beltran, M., Puig, I., Peña, C., García, J., Alvarez, A., Peña, R., et al. (2008). A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes & Development*, 22 (6), 756-69.

Ben-Shachar, S., Chahrour, M., Thaller, C., Shaw, C., & Zoghbi, H. (2009). Mouse models of MeCP2 disorders share gene expression changes in the cerebellum and hypothalamus. *Human molecular genetics*, 18 (13), 2431-42.

Berger, J., Sansom, O., Clarke, A., & Bird, A. (2007). MBD2 is required for correct spatial gene expression in the gut. *Molecular and cellular biology*, 27 (11), 4049-57.

Berger, S., Kouzarides, T., Shiekhatar, R., & Shilatfard, A. (2009). An operational definition of epigenetics. *Genes & Development*, 23 (7), 781-3.

Billard, L.-M., Magdinier, F., Lenoir, G., Frappart, L., & Dante, R. (2002). MeCP2 and MBD2 expression during normal and pathological growth of the human mammary gland. *Oncogene*, 21 (17), 2704-12.

Bird, A. (2007). Perceptions of epigenetics. *Nature*, 447 (7143), 396-8.

Blackwell, T., Huang, J., Ma, A., Kretzner, L., Alt, F., Eisenman, R., et al. (1993). Binding of myc proteins to canonical and noncanonical DNA sequences. *Molecular and Cellular Biology*, 13 (9), 5216-24.

Borgel, J., Guibert, S., Li, Y., Chiba, H., Schübeler, D., Sasaki, H., et al. (2010). Targets and dynamics of promoter DNA methylation during early mouse development. *Nature Genetics*, 42 (12), 1093-100.

Bostick, M., Kim, J., Estève, P.-O., Clark, A., Pradhan, S., & Jacobsen, S. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science (New York, NY)*, 317 (5845), 1760-4.

Bostick, M., Kim, J., Estève, P.-O., Clark, A., Pradhan, S., & Jacobsen, S. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science (New York, NY)*, 317 (5845), 1760-4.

Bourc'his, D., & Bestor, T. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature*, 431 (7004), 96-9.

Bourc'his, D., Xu, G., Lin, C., Bollman, B., & Bestor, T. (2001). Dnmt3L and the establishment of maternal genomic imprints. *Science (New York, NY)*, 294 (5551), 2536-9.

Bouzinba-Segard, H., Guais, A., & Francastel, C. (2006). Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. *Proceedings of the National Academy of Sciences of the United States of America*, 103 (23), 8709-14.

Bronner, C., Hopfner, R., & Mousli, M. (2002). Transcriptional regulation of the human topoisomerase IIalpha gene. *Anticancer research*, 22 (2A), 605-12.

Bronner, C., Trotzier, M.-A., Filhol, O., Cochet, C., Rochette-Egly, C., Schöller-Guinard, M., et al. (2004). The antiapoptotic protein ICBP90 is a target for protein kinase 2. *Annals of the New York Academy of Sciences*, 1030, 355-60.

Brown, C., Ballabio, A., Rupert, J., Lafreniere, R., Grompe, M., Tonlorenzi, R., et al. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature*, 349 (6304), 38-44.

Butler, M. (2002). Imprinting disorders: non-Mendelian mechanisms affecting growth. *Journal of pediatric endocrinology & metabolism : JPEM*, 15 Suppl 5, 1279-88.

Campanero, M., Armstrong, M., & Flemington, E. (2000). CpG methylation as a mechanism for the regulation of E2F activity. *Proceedings of the National Academy of Sciences of the United States of America*, 97 (12), 6481-6.

Carlson, L., Page, A., & Bestor, T. (1992). Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting. *Genes & Development*, 6 (12B), 2536-41.

Cedar, H., & Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nature reviews Genetics*, 10 (5), 295-304.

Chahrour, M., Jung, S., Shaw, C., Zhou, X., Wong, S., Qin, J., et al. (2008). MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science (New York, NY)* , 320 (5880), 1224-9.

Chang, B., Chen, Y., Zhao, Y., & Bruick, R. (2007). JMJD6 is a histone arginine demethylase. *Science (New York, NY)* , 318 (5849), 444-7.

Chao, H.-T., Chen, H., Samaco, R., Xue, M., Chahrour, M., Yoo, J., et al. (2010). Dysfunction in GABA signalling mediates autism-like stereotypies and Rett syndrome phenotypes. *Nature* , 468 (7321), 263-9.

Chatagnon, A., Ballestar, E., Esteller, M., & Dante, R. (2010). A role for methyl-CpG binding domain protein 2 in the modulation of the estrogen response of pS2/TFF1 gene. *PLoS ONE* , 5 (3), e9665.

Chatagnon, A., Bougel, S., Perriaud, L., Lachuer, J., Benhattar, J., & Dante, R. (2009). Specific association between the methyl-CpG-binding domain protein 2 and the hypermethylated region of the human telomerase reverse transcriptase promoter in cancer cells. *Carcinogenesis* , 30 (1), 28-34.

Chaumeil, J., Le Baccon, P., Wutz, A., & Heard, E. (2006). A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes & Development* , 20 (16), 2223-37.

Chedin, F., Lieber, M., & Hsieh, C.-L. (2002). The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proceedings of the National Academy of Sciences of the United States of America* , 99 (26), 16916-21.

Chen, R., Pettersson, U., Beard, C., Jackson-Grusby, L., & Jaenisch, R. (1998). DNA hypomethylation leads to elevated mutation rates. *Nature* , 395 (6697), 89-93.

Chen, W., Chang, Q., Lin, Y., Meissner, A., West, A., Griffith, E., et al. (2003). Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science (New York, NY)* , 302 (5646), 885-9.

Chodavarapu, R., Feng, S., Bernatavichute, Y., Chen, P.-Y., Stroud, H., Yu, Y., et al. (2010). Relationship between nucleosome positioning and DNA methylation. *Nature* , 466 (7304), 388-92.

Choo, K. (1997). Centromere DNA dynamics: latent centromeres and neocentromere formation. *American journal of human genetics* , 61 (6), 1225-33.

Chuang, L., Ian, H., Koh, T., Ng, H., Xu, G., & Li, B. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science (New York, NY)* , 277 (5334), 1996-2000.

Clouaire, T., de Las Heras, J., Merusi, C., & Stancheva, I. (2010). Recruitment of MBD1 to target genes requires sequence-specific interaction of the MBD domain with methylated DNA. *Nucleic Acids Research* , 38 (14), 4620-34.

Crawford, D., Acuña, J., & Sherman, S. (2001). FMR1 and the fragile X syndrome: human genome epidemiology review. *Genetics in medicine : official journal of the American College of Medical Genetics* , 3 (5), 359-71.

Csankovszki, G., Nagy, A., & Jaenisch, R. (2001). Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation. *The Journal of cell biology* , 153 (4), 773-84.

Damelin, M., & Bestor, T. (2007). Biological functions of DNA methyltransferase 1 require its methyltransferase activity. *Molecular and Cellular Biology* , 27 (11), 3891-9.

Daniel, J., Spring, C., Crawford, H., Reynolds, A., & Baig, A. (2002). The p120(ctn)-binding partner Kaiso is a bi-modal DNA-binding protein that recognizes both a sequence-specific consensus and methylated CpG dinucleotides. *Nucleic Acids Research* , 30 (13), 2911-9.

Danner, S., & Belasco, J. (2001). T7 phage display: a novel genetic selection system for cloning RNA-binding proteins from cDNA libraries. *Proceedings of the National Academy of Sciences of the United States of America* , 98 (23), 12954-9.

Dante, R., Dante-Paire, J., Rigal, D., & Roizès, G. (1992). Methylation patterns of long interspersed repeated DNA and alphoid repetitive DNA from human cell lines and tumors. *Anticancer research* , 12 (2), 559-63.

de Greef, J., Wang, J., Balog, J., den Dunnen, J., Frants, R., Straasheijm, K., et al. (2011). Mutations in ZBTB24 Are Associated with Immunodeficiency, Centromeric Instability, and Facial Anomalies Syndrome Type 2. *American journal of human genetics* , 88 (6), 796-804.

D'Herelle, F. (2007). On an invisible microbe antagonistic toward dysenteric bacilli: brief note by Mr. F. D'Herelle, presented by Mr. Roux. 1917. *Research in microbiology* , 158 (7), 553-4.

Dodge, J., Ramsahoye, B., Wo, Z., Okano, M., & Li, E. (2002). De novo methylation of MMLV provirus in embryonic stem cells: CpG versus non-CpG methylation. *Gene* , 289 (1-2), 41-8.

Doi, A., Park, I.-H., Wen, B., Murakami, P., Aryee, M., Irizarry, R., et al. (2009). Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nature Genetics* , 41 (12), 1350-3.

Drané, P., Ouvarhni, K., Depaux, A., Shuaib, M., & Hamiche, A. (2010). The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes & Development* , 24 (12), 1253-65.

Ego, T., Tanaka, Y., & Shimotohno, K. (2005). Interaction of HTLV-1 Tax and methyl-CpG-binding domain 2 positively regulates the gene expression from the hypermethylated LTR. *Oncogene* , 24 (11), 1914-23.

Ehrlich, M. (2003). The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. *Clinical immunology (Orlando, Fla)* , 109 (1), 17-28.

Elliott, M., & Maher, E. (1994). Beckwith-Wiedemann syndrome. *Journal of medical genetics* , 31 (7), 560-4.

Ernst, J., & Kellis, M. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nature biotechnology* , 28 (8), 817-25.

Ernst, J., Kheradpour, P., Mikkelsen, T., Shores, N., Ward, L., Epstein, C., et al. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* , 473 (7345), 43-9.

Espada, J., Ballestar, E., Santoro, R., Fraga, M., Villar-Garea, A., Németh, A., et al. (2007). Epigenetic disruption of ribosomal RNA genes and nucleolar architecture in DNA methyltransferase 1 (Dnmt1) deficient cells. *Nucleic Acids Research* , 35 (7), 2191-8.

Esteller, M. (2007). Epigenetic gene silencing in cancer: the DNA hypermethylome. *Human Molecular Genetics* , 16 Spec No 1, R50-9.

Esteller, M. (2008). Epigenetics in cancer. *The New England journal of medicine* , 358 (11), 1148-59.

Estève, P.-O., Chang, Y., Samaranayake, M., Upadhyay, A., Horton, J., Feehery, G., et al. (2011). A methylation and phosphorylation switch between an adjacent lysine and serine determines human DNMT1 stability. *Nature Structural & Molecular Biology* , 18 (1), 42-8.

Estève, P.-O., Chin, H., Benner, J., Feehery, G., Samaranayake, M., Horwitz, G., et al. (2009). Regulation of DNMT1 stability through SET7-mediated lysine methylation in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* , 106 (13), 5076-81.

Estève, P.-O., Chin, H., Smallwood, A., Feehery, G., Gangisetty, O., Karpf, A., et al. (2006). Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes & Development* , 20 (22), 3089-103.

Fatemi, M., Hermann, A., Gowher, H., & Jeltsch, A. (2002). Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. *Federation of European of Biochemical Societies Journal / FEBS Journal* , 269 (20), 4981-4.

Feng, S., Cokus, S., Zhang, X., Chen, P.-Y., Bostick, M., Goll, M., et al. (2010). Conservation and divergence of methylation patterning in plants and animals. *Proceedings of the National Academy of Sciences of the United States of America* , 107 (19), 8689-94.

Feng, S.-Y., Ota, K., & Ito, T. (2010). A yeast one-hybrid system to screen for methylated DNA-binding proteins. *Nucleic Acids Research* , 38 (20), e189.

Feng, S.-Y., Ota, K., Yamada, Y., Sawabu, N., & Ito, T. (2004). A yeast one-hybrid system to detect methylation-dependent DNA-protein interactions. *Biochemical and Biophysical Research Communications* , 313 (4), 922-5.

Ferri, F., Bouzinba-Segard, H., Velasco, G., Hubé, F., & Francastel, C. (2009). Non-coding murine centromeric transcripts associate with and potentiate Aurora B kinase. *Nucleic Acids Research* , 37 (15), 5071-80.

Ficz, G., Branco, M., Seisenberger, S., Santos, F., Krueger, F., Hore, T., et al. (2011). Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* , 473 (7347), 398-402.

Filion, G., Zhenilo, S., Salozhin, S., Yamada, D., Prokhortchouk, E., & Defossez, P.-A. (2006). A family of human zinc finger proteins that bind methylated DNA and repress transcription. *Molecular and cellular biology* , 26 (1), 169-81.

Fisher, D., Carr, C., Parent, L., & Sharp, P. (1991). TFEB has DNA-binding and oligomerization properties of a unique helix-loop-helix/leucine-zipper family. *Genes & Development* , 5 (12A), 2342-52.

Fournier, C., Goto, Y., Ballestar, E., Delaval, K., Hever, A., Esteller, M., et al. (2002). Allele-specific histone lysine methylation marks regulatory regions at imprinted mouse genes. *The EMBO journal* , 21 (23), 6560-70.

Frauer, C., Hoffmann, T., Bultmann, S., Casa, V., Cardoso, M., Antes, I., et al. (2011). Recognition of 5-Hydroxymethylcytosine by the Uhrf1 SRA Domain. *PloS one* , 6 (6), e21306.

Fujimori, A., Matsuda, Y., Takemoto, Y., Hashimoto, Y., Kubo, E., Araki, R., et al. (1998). Cloning and mapping of Np95 gene which encodes a novel nuclear protein associated with cell proliferation. *Mammalian genome : official journal of the International Mammalian Genome Society* , 9 (12), 1032-5.

Fujita, N., Watanabe, S., Ichimura, T., Tsuruzoe, S., Shinkai, Y., Tachibana, M., et al. (2003). Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression. *The Journal of biological chemistry* , 278 (26), 24132-24138.

Fuks, F., Burgers, W., Brehm, A., Hughes-Davies, L., & Kouzarides, T. (2000). DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nature Genetics* , 24 (1), 88-91.

Fuks, F., Hurd, P., Deplus, R., & Kouzarides, T. (2003). The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Research* , 31 (9), 2305-12.

Futscher, B., Oshiro, M., Wozniak, R., Holtan, N., Hanigan, C., Duan, H., et al. (2002). Role for DNA methylation in the control of cell type specific maspin expression. *Nature Genetics* , 31 (2), 175-9.

Garzon, R., Heaphy, C., Havelange, V., Fabbri, M., Volinia, S., Tsao, T., et al. (2009). MicroRNA 29b functions in acute myeloid leukemia. *Blood* , 114 (26), 5331-41.

Gaudet, F., Hodgson, J., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J., et al. (2003). Induction of tumors in mice by genomic hypomethylation. *Science (New York, NY)* , 300 (5618), 489-92.

Gazin, C., Wajapeyee, N., Gobeil, S., Virbasius, C.-M., & Green, M. (2007). An elaborate pathway required for Ras-mediated epigenetic silencing. *Nature* , 449 (7165), 1073-7.

Ghosh, R., Horowitz-Scherer, R., Nikitina, T., Shlyakhtenko, L., & Woodcock, C. (2010). MeCP2 binds cooperatively to its substrate and competes with histone H1 for chromatin binding sites. *Molecular and cellular biology* , 30 (19), 4656-70.

Giese, K., Amsterdam, A., & Grosschedl, R. (1991). DNA-binding properties of the HMG domain of the lymphoid-specific transcriptional regulator LEF-1. *Genes & Development*, 5 (12B), 2567-78.

Glass, J., Fazzari, M., Ferguson-Smith, A., & Greal, J. (2009). CG dinucleotide periodicities recognized by the Dnmt3a-Dnmt3L complex are distinctive at retroelements and imprinted domains. *Mammalian genome : official journal of the International Mammalian Genome Society*, 20 (9-10), 633-43.

Globisch, D., Münzel, M., Müller, M., Michalakakis, S., Wagner, M., Koch, S., et al. (2010). Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PloS one*, 5 (12), e15367.

Goelz, S., Vogelstein, B., Hamilton, S., & Feinberg, A. (1985). Hypomethylation of DNA from benign and malignant human colon neoplasms. *Science (New York, NY)*, 228 (4696), 187-90.

Goldstone, A. (2004). Prader-Willi syndrome: advances in genetics, pathophysiology and treatment. *Trends in endocrinology and metabolism: TEM*, 15 (1), 12-20.

Goll, M., Kirpekar, F., Maggert, K., Yoder, J., Hsieh, C.-L., Zhang, X., et al. (2006). Methylation of tRNA<sup>Asp</sup> by the DNA methyltransferase homolog Dnmt2. *Science (New York, NY)*, 311 (5759), 395-8.

Gonzalo, S., Jaco, I., Fraga, M., Chen, T., Li, E., Esteller, M., et al. (2006). DNA methyltransferases control telomere length and telomere recombination in mammalian cells. *Nature Cell Biology*, 8 (4), 416-24.

Gowher, H., Liebert, K., Hermann, A., Xu, G., & Jeltsch, A. (2005). Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. *The Journal of biological chemistry*, 280 (14), 13341-8.

Goyal, R., Rathert, P., Laser, H., Gowher, H., & Jeltsch, A. (2007). Phosphorylation of serine-515 activates the Mammalian maintenance methyltransferase Dnmt1. *Epigenetics : official journal of the DNA Methylation Society*, 2 (3), 155-60.

Gu, P., Xu, X., Le Menuet, D., Chung, A.-K., & Cooney, A. (2011). Differential Recruitment of Methyl CpG-Binding Domain Factors and DNA Methyltransferases by the Orphan Receptor Germ Cell Nuclear Factor Initiates the Repression and Silencing of Oct4. *Stem cells (Dayton, Ohio)*, 29 (7), 1041-51.

Guo, H., Lin, Y., Zhang, H., Liu, J., Zhang, N., Li, Y., et al. (2007). Tissue factor pathway inhibitor-2 was repressed by CpG hypermethylation through inhibition of KLF6 binding in highly invasive breast cancer cells. *BMC molecular biology*, 8, 110.

Guo, J., Su, Y., Zhong, C., Ming, G.-l., & Song, H. (2011). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell*, 145 (3), 423-34.

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., et al. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mechanisms of development*, 117 (1-2), 15-23.

Hansen, R., Stöger, R., Wijmenga, C., Stanek, A., Canfield, T., Luo, P., et al. (2000). Escape from gene silencing in ICF syndrome: evidence for advanced replication time as a major determinant. *Human molecular genetics*, 9 (18), 2575-87.

Hata, K., Okano, M., Lei, H., & Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development (Cambridge, England)*, 129 (8), 1983-93.

Hendrich, B., & Bird, A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Molecular and cellular biology*, 18 (11), 6538-47.

Hendrich, B., & Tweedie, S. (2003). The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *Trends in genetics : TIG*, 19 (5), 269-77.

Hendrich, B., Guy, J., Ramsahoye, B., Wilson, V., & Bird, A. (2001). Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. *Genes & development*, 15 (6), 710-23.

- Hendrich, B., Hardeland, U., Ng, H., Jiricny, J., & Bird, A. (1999). The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature* , 401 (6750), 301-4.
- Hermann, A., Schmitt, S., & Jeltsch, A. (2003). The human Dnmt2 has residual DNA- (cytosine-C5) methyltransferase activity. *The Journal of biological chemistry* , 278 (34), 31717-21.
- Hervouet, E., Lalier, L., Debieu, E., Cheray, M., Geairon, A., Rogniaux, H., et al. (2010). Disruption of Dnmt1/PCNA/UHRF1 interactions promotes tumorigenesis from human and mice glial cells. *PloS one* , 5 (6), e11333.
- Hervouet, E., Vallette, F., & Cartron, P.-F. (2010 йил 29-Jan). Dnmt1/transcription factor interactions: an alternative mechanism of DNA methylation inheritance. *Genes and Cancer for peer review* , 1-27.
- Ho, K., McNae, I., Schmiedeberg, L., Klose, R., Bird, A., & Walkinshaw, M. (2008). MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Molecular Cell* , 29 (4), 525-31.
- Holley, R., Apgar, J., & Everett, G. (1965). STRUCTURE OF A RIBONUCLEIC ACID. *Science (New York)* .
- Holley, R., Everett, G., & Madison, J. (1965). NUCLEOTIDE SEQUENCES IN THE YEAST ALANINE TRANSFER RIBONUCLEIC ACID. *The Journal of biological ...* .
- Hopfner, R., Mousli, M., Jeltsch, J., Voulgaris, A., Lutz, Y., Marin, C., et al. (2000). ICBP90, a novel human CCAAT binding protein, involved in the regulation of topoisomerase IIalpha expression. *Cancer research* , 60 (1), 121-8.
- Hopfner, R., Mousli, M., Oudet, P., & Bronner, C. (2002). Overexpression of ICBP90, a novel CCAAT-binding protein, overcomes cell contact inhibition by forcing topoisomerase II alpha expression. *Anticancer research* , 22 (6A), 3165-70.
- Hutchins, A., Mullen, A., Lee, H., Sykes, K., High, F., Hendrich, B., et al. (2002). Gene silencing quantitatively controls the function of a developmental trans-activator. *Molecular cell* , 10 (1), 81-91.
- Iguchi-Arigo, S., & Schaffner, W. (1989). CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. *Genes & Development* , 3 (5), 612-9.
- Illingworth, R., Gruenewald-Schneider, U., Webb, S., Kerr, A., James, K., Turner, D., et al. (2010). Orphan CpG islands identify numerous conserved promoters in the mammalian genome. *PLoS genetics* , 6 (9).
- Illingworth, R., Kerr, A., Desousa, D., Jørgensen, H., Ellis, P., Stalker, J., et al. (2008). A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biology* , 6 (1), e22.
- Irizarry, R., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., et al. (2009). The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nature Genetics* , 41 (2), 178-86.
- Ito, Y., Koessler, T., Ibrahim, A., Rai, S., Vowler, S., Abu-Amero, S., et al. (2008). Somatically acquired hypomethylation of IGF2 in breast and colorectal cancer. *Human molecular genetics* , 17 (17), 2633-43.
- Jørgensen, H., Ben-Porath, I., & Bird, A. (2004). Mbd1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains. *Molecular and Cellular Biology* , 24 (8), 3387-95.
- Jacob, F., & Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins†. *Journal of Molecular Biology* , 3 (3), 318-356.
- Jacobs, S., & Khorasanizadeh, S. (2002). Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science (New York, NY)* , 295 (5562), 2080-2083.
- Jair, K.-W., Bachman, K., Suzuki, H., Ting, A., Rhee, I., Yen, R.-W., et al. (2006). De novo CpG island methylation in human cancer cells. *Cancer research* , 66 (2), 682-92.

Jeanblanc, M., Mousli, M., Hopfner, R., Bathami, K., Martinet, N., Abbady, A.-Q., et al. (2005). The retinoblastoma gene and its product are targeted by ICBP90: a key mechanism in the G1/S transition during the cell cycle. *Oncogene* , 24 (49), 7337-45.

Jia, D., Jurkowska, R., Zhang, X., Jeltsch, A., & Cheng, X. (2007). Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* , 449 (7159), 248-51.

Jiang, G., Yang, F., van Overveld, P., Vedanarayanan, V., van der Maarel, S., & Ehrlich, M. (2003). Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. *Human molecular genetics* , 12 (22), 2909-21.

Jin, S.-G., Wu, X., Li, A., & Pfeifer, G. (2011). Genomic mapping of 5-hydroxymethylcytosine in the human brain. *Nucleic Acids Research* , 39 (12), 5015-24.

Kaneda, M., Hirasawa, R., Chiba, H., Okano, M., Li, E., & Sasaki, H. (2010 йил 1-Feb). Genetic evidence for Dnmt3a-dependent imprinting during oocyte growth obtained by conditional knockout with Zp3-Cre and complete exclusion of Dnmt3b by chimera formation. *Genes to cells : devoted to molecular & cellular mechanisms* .

Karagianni, P., Amazit, L., Qin, J., & Wong, J. (2008). ICBP90, a novel methyl K9 H3 binding protein linking protein ubiquitination with heterochromatin formation. *Molecular and cellular biology* , 28 (2), 705-17.

Kareta, M., Botello, Z., Ennis, J., Chou, C., & Chédin, F. (2006). Reconstitution and mechanism of the stimulation of de novo methylation by human DNMT3L. *The Journal of biological chemistry* , 281 (36), 25893-902.

Kim, J., Estève, P.-O., Jacobsen, S., & Pradhan, S. (2009). UHRF1 binds G9a and participates in p21 transcriptional regulation in mammalian cells. *Nucleic Acids Research* , 37 (2), 493-505.

Kim, S., Park, J.-I., Spring, C., Sater, A., Ji, H., Otchere, A., et al. (2004). Non-canonical Wnt signals are modulated by the Kaiso transcriptional repressor and p120-catenin. *Nature Cell Biology* , 6 (12), 1212-20.

Kinoshita, K., Kikuchi, Y., Sasakura, Y., Suzuki, M., Fujii-Kuriyama, Y., & Sogawa, K. (2004). Altered DNA binding specificity of Arnt by selection of partner bHLH-PAS proteins. *Nucleic Acids Research* , 32 (10), 3169-79.

Klose, R., Sarraf, S., Schmiedeberg, L., McDermott, S., Stancheva, I., & Bird, A. (2005). DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. *Molecular Cell* , 19 (5), 667-78.

Kondo, E., Gu, Z., Horii, A., & Fukushige, S. (2005). The thymine DNA glycosylase MBD4 represses transcription and is associated with methylated p16(INK4a) and hMLH1 genes. *Molecular and Cellular Biology* , 25 (11), 4388-96.

Kornberg, A., Bertsch, L., Jackson, J., Khorana, H. (1964). Enzymatic synthesis of deoxyribonucleic acid, XVI. Oligonucleotides as templates and the mechanism of their replication. *Proceedings of the National Academy of Sciences of the United States of America* , 51, 315-323.

Kransdorf, E., Wang, S., Zhu, S., Langston, T., Rupon, J., & Ginder, G. (2006). MBD2 is a critical component of a methyl cytosine-binding protein complex isolated from primary erythroid cells. *Blood* , 108 (8), 2836-45.

Kriaucionis, S., & Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science (New York, NY)* , 324 (5929), 929-30.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., et al. (2008). DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes & Development* , 22 (7), 908-17.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., & Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* , 410 (6824), 116-120.

Lane, N., Dean, W., & Reik, W. (2003). Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* .

Laurent, L., Wong, E., Li, G., Huynh, T., Tsigos, A., Ong, C., et al. (2010). Dynamic changes in the human methylome during differentiation. *Genome research* , 20 (3), 320-31.

Lawrence, S. (2007 йил 27-Mar). Billion dollar babies—biotech drugs as blockbusters. 1-3.

Lee, B., & Muller, M. (2009). SUMOylation enhances DNA methyltransferase 1 activity. *The Biochemical journal* , 421 (3), 449-61.

Lehnertz, B., Ueda, Y., Derijck, A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., et al. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Current biology : CB* , 13 (14), 1192-200.

Lei, H., Oh, S., Okano, M., Jüttermann, R., Goss, K., Jaenisch, R., et al. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development (Cambridge, England)* , 122 (10), 3195-205.

Lewis, J., Meehan, R., Henzel, W., Maurer-Fogy, I., Jeppesen, P., Klein, F., et al. (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* , 69 (6), 905-14.

Li, E., Beard, C., & Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. *Nature* , 366 (6453), 362-5.

Li, E., Bestor, T., & Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* , 69 (6), 915-26.

Li, X., Barkho, B., Luo, Y., Smrt, R., Santistevan, N., Liu, C., et al. (2008). Epigenetic regulation of the stem cell mitogen Fgf-2 by Mbd1 in adult neural stem/progenitor cells. *The Journal of biological chemistry* , 283 (41), 27644-52.

Li, Y., Mori, T., Hata, H., Homma, Y., & Kochi, H. (2004). NIRF induces G1 arrest and associates with Cdk2. *Biochemical and biophysical research communications* , 319 (2), 464-8.

Lin, C.-H., Hansen, S., Wang, Z., Storm, D., Tapscott, S., & Olson, J. (2005). The dosage of the neuroD2 transcription factor regulates amygdala development and emotional learning. *Proceedings of the National Academy of Sciences of the United States of America* , 102 (41), 14877-82.

Lister, R., Pelizzola, M., Dowen, R., Hawkins, R., Hon, G., Tonti-Filippini, J., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* , 462 (7271), 315-22.

Liu, C., Teng, Z.-Q., Santistevan, N., Szulwach, K., Guo, W., Jin, P., et al. (2010). Epigenetic regulation of miR-184 by MBD1 governs neural stem cell proliferation and differentiation. *Cell stem cell* , 6 (5), 433-44.

Lossie, A., Whitney, M., Amidon, D., Dong, H., Chen, P., Theriaque, D., et al. (2001). Distinct phenotypes distinguish the molecular classes of Angelman syndrome. *Journal of medical genetics* , 38 (12), 834-45.

Luco, R., Pan, Q., Tominaga, K., Blencowe, B., Pereira-Smith, O., & Misteli, T. (2010). Regulation of alternative splicing by histone modifications. *Science (New York, NY)* , 327 (5968), 996-1000.

Luger, K., Mäder, A., Richmond, R., Sargent, D., & Richmond, T. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* , 389 (6648), 251-60.

Lujambio, A., Roperio, S., Ballestar, E., Fraga, M., Cerrato, C., Setién, F., et al. (2007). Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer research* , 67 (4), 1424-9.

Lyko, F., Ramsahoye, B., & Jaenisch, R. (2000). DNA methylation in *Drosophila melanogaster*. *Nature* , 408 (6812), 538-40.

Lyon, M. (1961). Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* (190), 372-373.

Maatouk, D., Kellam, L., Mann, M., Lei, H., Li, E., Bartolomei, M., et al. (2006). DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages. *Development* , 133 (17), 3411-8.



Magdinier, F., Billard, L., Wittmann, G., Frappart, L., Benchaïb, M., Lenoir, G., et al. (2000). Regional methylation of the 5' end CpG island of BRCA1 is associated with reduced gene expression in human somatic cells. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology* , 14 (11), 1585-94.

Magdinier, F., Billard, L., Wittmann, G., Frappart, L., Benchaïb, M., Lenoir, G., et al. (2000). Regional methylation of the 5' end CpG island of BRCA1 is associated with reduced gene expression in human somatic cells. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology* , 14 (11), 1585-1594.

Magdinier, F., D'Estaing, S., Peinado, C., Demirci, B., Berthet, C., Guérin, J., et al. (2002). Epigenetic marks at BRCA1 and p53 coding sequences in early human embryogenesis. *Molecular human reproduction* , 8 (7), 630-5.

Martín Caballero, I., Hansen, J., Leaford, D., Pollard, S., & Hendrich, B. (2009). The methyl-CpG binding proteins Mecp2, Mbd2 and Kaiso are dispensable for mouse embryogenesis, but play a redundant function in neural differentiation. *PLoS ONE* , 4 (1), e4315.

Martin, V., Ribieras, S., Song-Wang, X., Lasne, Y., Frappart, L., Rio, M., et al. (1997). Involvement of DNA methylation in the control of the expression of an estrogen-induced breast-cancer-associated protein (pS2) in human breast cancers. *Journal of Cellular Biochemistry* , 65 (1), 95-106.

Martin, V., Ribieras, S., Song-Wang, X., Rio, M., & Dante, R. (1995). Genomic sequencing indicates a correlation between DNA hypomethylation in the 5' region of the pS2 gene and its expression in human breast cancer cell lines. *Gene* , 157 (1-2), 261-4.

Martin, V., Ribieras, S., Song-Wang, X., Rio, M., & Dante, R. (1995). Genomic sequencing indicates a correlation between DNA hypomethylation in the 5' region of the pS2 gene and its expression in human breast cancer cell lines. *Gene* , 157 (1-2), 261-264.

Massari, M., & Murre, C. (2000 йил 10-Dec). Helix-Loop-Helix Proteins: Regulators of Transcription in Eucaryotic Organisms. *Molecular and Cellular Biology* , 1-12.

Matsumura, S., Persson, L., Wong, L., & Wilson, A. (2010). The latency-associated nuclear antigen interacts with MeCP2 and nucleosomes through separate domains. *Journal of Virology* , 84 (5), 2318-30.

MATTHAEI, J., JONES, O., MARTIN, R., & NIRENBERG, M. (1962). Characteristics and composition of RNA coding units. *Proceedings of the National Academy of Sciences of the United States of America* , 48, 666-677.

Maunakea, A., Nagarajan, R., Bilenky, M., Ballinger, T., D'Souza, C., Fouse, S., et al. (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* , 466 (7303), 253-7.

Mayer, W., Niveleau, A., Walter, J., Fundele, R., & Haaf, T. (2000). Demethylation of the zygotic paternal genome. *Nature* , 403 (6769), 501-2.

McGrath, J., & Solter, D. (1983). Nuclear transplantation in the mouse embryo by microsurgery and cell fusion. *Science (New York, NY)* , 220 (4603), 1300-2.

Meehan, R., Lewis, J., McKay, S., Kleiner, E., & Bird, A. (1989). Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* , 58 (3), 499-507.

Meilinger, D., Fellinger, K., Bultmann, S., Rothbauer, U., Bonapace, I., Klinkert, W., et al. (2009). Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells. *EMBO reports* , 10 (11), 1259-64.

Meissner, A., Mikkelsen, T., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., et al. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* , 454 (7205), 766-70.

Meselson, M., & Stahl, F. (1958). THE REPLICATION OF DNA IN ESCHERICHIA COLI. *Proceedings of the National Academy of Sciences of the United States of America* , 44 (7), 671-682.

Millar, C., Guy, J., Sansom, O., Selfridge, J., MacDougall, E., Hendrich, B., et al. (2002). Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. *Science (New York, NY)* , 297 (5580), 403-5.

Mills, R., Bennett, E., Iskow, R., & Devine, S. (2007). Which transposable elements are active in the human genome? *Trends in genetics : TIG* , 23 (4), 183-91.

Moazed, D. (2009). Small RNAs in transcriptional gene silencing and genome defence. *Nature* , 457 (7228), 413-20.

Mohammad, F., Mondal, T., Guseva, N., Pandey, G., & Kanduri, C. (2010). Kcnq1ot1 noncoding RNA mediates transcriptional gene silencing by interacting with Dnmt1. *Development (Cambridge, England)* , 137 (15), 2493-9.

Morgan, T. (1910). Sex limited inheritance in *Drosophila*. *Science (New York, NY)* .

Mori, T., Li, Y., Hata, H., & Kochi, H. (2004). NIRF is a ubiquitin ligase that is capable of ubiquitinating PCNP, a PEST-containing nuclear protein. *FEBS letters* , 557 (1-3), 209-14.

Mori, T., Li, Y., Hata, H., Ono, K., & Kochi, H. (2002). NIRF, a novel RING finger protein, is involved in cell-cycle regulation. *Biochemical and biophysical research communications* , 296 (3), 530-6.

Morison, I., Ramsay, J., & Spencer, H. (2005). A census of mammalian imprinting. *Trends in genetics : TIG* , 21 (8), 457-65.

Mortusewicz, O., Schermelleh, L., Walter, J., Cardoso, M., & Leonhardt, H. (2005). Recruitment of DNA methyltransferase I to DNA repair sites. *Proceedings of the National Academy of Sciences of the United States of America* , 102 (25), 8905-9.

Moulton, T., Crenshaw, T., Hao, Y., Moosikasuwan, J., Lin, N., Dembitzer, F., et al. (1994). Epigenetic lesions at the H19 locus in Wilms' tumour patients. *Nature Genetics* , 7 (3), 440-7.

Mousli, M., Hopfner, R., Abbady, A.-Q., Monté, D., Jeanblanc, M., Oudet, P., et al. (2003). ICBP90 belongs to a new family of proteins with an expression that is deregulated in cancer cells. *British Journal of Cancer* , 89 (1), 120-7.

Muto, M., Kanari, Y., Kubo, E., Takabe, T., Kurihara, T., Fujimori, A., et al. (2002). Targeted disruption of Np95 gene renders murine embryonic stem cells hypersensitive to DNA damaging agents and DNA replication blocks. *The Journal of biological chemistry* , 277 (37), 34549-55.

Nady, N., Lemak, A., Walker, J., Avvakumov, G., Kareta, M., Achour, M., et al. (2011 йил 13-Apr). Recognition of multivalent histone states associated with heterochromatin by UHRF1. *The Journal of biological chemistry* .

Nagano, T., Mitchell, J., Sanz, L., Pauler, F., Ferguson-Smith, A., Feil, R., et al. (2008). The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science (New York, NY)* , 322 (5908), 1717-20.

Nan, X., Hou, J., Maclean, A., Nasir, J., Lafuente, M., Shu, X., et al. (2007). Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. *Proceedings of the National Academy of Sciences of the United States of America* , 104 (8), 2709-14.

Nan, X., Meehan, R., & Bird, A. (1993). Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Research* , 21 (21), 4886-92.

Nan, X., Ng, H., Johnson, C., Laherty, C., Turner, B., Eisenman, R., et al. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* , 393 (6683), 386-9.

Nan, X., Tate, P., Li, E., & Bird, A. (1996). DNA methylation specifies chromosomal localization of MeCP2. *Molecular and Cellular Biology* , 16 (1), 414-21.

Ng, H., Zhang, Y., Hendrich, B., Johnson, C., Turner, B., Erdjument-Bromage, H., et al. (1999). MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nature genetics* , 23 (1), 58-61.

Nicholls, R., & Knepper, J. (2001). Genome organization, function, and imprinting in Prader-Willi and Angelman syndromes. *Annual review of genomics and human genetics* , 2, 153-75.

- Nielsen, A., Oulad-Abdelghani, M., Ortiz, J., Remboutsika, E., Chambon, P., & Losson, R. (2001). Heterochromatin formation in mammalian cells: interaction between histones and HP1 proteins. *Molecular Cell*, 7 (4), 729-739.
- Nikitina, T., Ghosh, R., Horowitz-Scherer, R., Hansen, J., Grigoryev, S., & Woodcock, C. (2007). MeCP2-chromatin interactions include the formation of chromatosome-like structures and are altered in mutations causing Rett syndrome. *The Journal of biological chemistry*, 282 (38), 28237-45.
- Nowak, S., & Corces, V. (2004). Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends in genetics : TIG*, 20 (4), 214-20.
- Oberlé, I., Rousseau, F., Heitz, D., Kretz, C., Devys, D., Hanauer, A., et al. (1991). Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science (New York, NY)*, 252 (5010), 1097-102.
- Oda, M., Yamagiwa, A., Yamamoto, S., Nakayama, T., Tsumura, A., Sasaki, H., et al. (2006). DNA methylation regulates long-range gene silencing of an X-linked homeobox gene cluster in a lineage-specific manner. *Genes & Development*, 20 (24), 3382-94.
- Ohki, I., Shimotake, N., Fujita, N., Jee, J., Ikegami, T., Nakao, M., et al. (2001). Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. *Cell*, 105 (4), 487-97.
- OHNO, S., KAPLAN, W., & KINOSITA, R. (1959). Formation of the sex chromatin by a single X-chromosome in liver cells of *Rattus norvegicus*. *Experimental cell research*, 18, 415-418.
- Okano, M., Xie, S., & Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature genetics*, 19 (3), 219-20.
- Olek, A., & Walter, J. (1997). The pre-implantation ontogeny of the H19 methylation imprint. *Nature Genetics*.
- Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D., Steen, H., Pandey, A., et al. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & cellular proteomics : MCP*, 1 (5), 376-86.
- Ooi, S., & Bestor, T. (2008). The colorful history of active DNA demethylation. *Cell*, 133 (7), 1145-8.
- Ooi, S., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., et al. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature*, 448 (7154), 714-7.
- Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., et al. (2000). Active demethylation of the paternal genome in the mouse zygote. *Current biology : CB*, 10 (8), 475-8.
- Pandey, R., Mondal, T., Mohammad, F., Enroth, S., Redrup, L., Komorowski, J., et al. (2008). Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Molecular Cell*, 32 (2), 232-46.
- Papait, R., Pistore, C., Grazini, U., Babbio, F., Cogliati, S., Pecoraro, D., et al. (2008). The PHD domain of Np95 (mUHRF1) is involved in large-scale reorganization of pericentromeric heterochromatin. *Molecular biology of the cell*, 19 (8), 3554-63.
- Papait, R., Pistore, C., Negri, D., Pecoraro, D., Cantarini, L., & Bonapace, I. (2007). Np95 is implicated in pericentromeric heterochromatin replication and in major satellite silencing. *Molecular biology of the cell*, 18 (3), 1098-106.
- Park, J.-I., Kim, S., Lyons, J., Ji, H., Nguyen, T., Cho, K., et al. (2005). Kaiso/p120-catenin and TCF/beta-catenin complexes coordinately regulate canonical Wnt gene targets. *Developmental cell*, 8 (6), 843-54.
- Penny, G., Kay, G., Sheardown, S., Rastan, S., & Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. *Nature*, 379 (6561), 131-7.

Perez-Torrado, R., Yamada, D., & Defossez, P.-A. (2006). Born to bind: the BTB protein-protein interaction domain. *BioEssays : news and reviews in molecular, cellular and developmental biology* , 28 (12), 1194-202.

Perrin, D., Ballestar, E., Fraga, M., Frappart, L., Esteller, M., Guerin, J.-F., et al. (2007). Specific hypermethylation of LINE-1 elements during abnormal overgrowth and differentiation of human placenta. *Oncogene* , 26 (17), 2518-24.

Petosa, C., Morand, P., Baudin, F., Moulin, M., Artero, J.-B., & Müller, C. (2006). Structural basis of lytic cycle activation by the Epstein-Barr virus ZEBRA protein. *Molecular Cell* , 21 (4), 565-72.

Phalke, S., Nickel, O., Walluscheck, D., Hortig, F., Onorati, M., & Reuter, G. (2009). Retrotransposon silencing and telomere integrity in somatic cells of *Drosophila* depends on the cytosine-5 methyltransferase DNMT2. *Nature Genetics* , 41 (6), 696-702.

Pichler, G., & Leonhardt, H. (2011 йил 9-May). Cooperative DNA and histone binding by Uhrf2 links the two major repressive epigenetic pathways. *Journal of Cellular Biochemistry* , 1-28.

Pichler, G., Wolf, P., Schmidt, C., Meilinger, D., Schneider, K., Frauer, C., et al. (2011 йил 19-May). Cooperative DNA and histone binding by Uhrf2 links the two major repressive epigenetic pathways. *Journal of Cellular Biochemistry* .

Plath, K., Fang, J., Mlynarczyk-Evans, S., Cao, R., Worringer, K., Wang, H., et al. (2003). Role of histone H3 lysine 27 methylation in X inactivation. *Science (New York, NY)* , 300 (5616), 131-5.

Popp, C., Dean, W., Feng, S., Cokus, S., Andrews, S., Pellegrini, M., et al. (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* , 463 (7284), 1101-5.

Prendergast, G., & Ziff, E. (1991). Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region. *Science (New York, NY)* , 251 (4990), 186-9.

Probst, A., & Almouzni, G. (2011). Heterochromatin establishment in the context of genome-wide epigenetic reprogramming. *Trends in Genetics* , 27 (5), 177-85.

Probst, A., Dunleavy, E., & Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. *Nature reviews Molecular cell biology* , 10 (3), 192-206.

Prokhortchouk, A., Hendrich, B., Jørgensen, H., Ruzov, A., Wilm, M., Georgiev, G., et al. (2001). The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. *Genes & Development* , 15 (13), 1613-8.

Prokhortchouk, A., Sansom, O., Selfridge, J., Caballero, I., Salozhin, S., Aithozhina, D., et al. (2006). Kaiso-deficient mice show resistance to intestinal cancer. *Molecular and Cellular Biology* , 26 (1), 199-208.

Rai, K., Huggins, I., James, S., Karpf, A., Jones, D., & Cairns, B. (2008). DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* , 135 (7), 1201-12.

Ramsahoye, B., Biniszkievicz, D., Lyko, F., Clark, V., Bird, A., & Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proceedings of the National Academy of Sciences of the United States of America* , 97 (10), 5237-42.

Reik, W., & Walter, J. (2001). Genomic imprinting: parental influence on the genome. *Nature reviews Genetics* , 2 (1), 21-32.

Reik, W., Dean, W., & Walter, J. (2001). Epigenetic reprogramming in mammalian development. *Science (New York, NY)* , 293 (5532), 1089-93.

Ribieras, S., Song-Wang, X., Martin, V., Lointier, P., Frappart, L., & Dante, R. (1994). Human breast and colon cancers exhibit alterations of DNA methylation patterns at several DNA segments on chromosomes 11p and 17p. *Journal of Cellular Biochemistry* , 56 (1), 86-96.

Rinn, J., Kertesz, M., Wang, J., Squazzo, S., Xu, X., Brugmann, S., et al. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*, 129 (7), 1311-23.

Robertson, K., Keyomarsi, K., Gonzales, F., Velicescu, M., & Jones, P. (2000). Differential mRNA expression of the human DNA methyltransferases (DNMTs) 1, 3a and 3b during the G(0)/G(1) to S phase transition in normal and tumor cells. *Nucleic acids research*, 28 (10), 2108-13.

Roloff, T., Ropers, H., & Nuber, U. (2003). Comparative study of methyl-CpG-binding domain proteins. *BMC genomics*, 4 (1), 1.

Rottach, A., Frauer, C., Pichler, G., Bonapace, I., Spada, F., & Leonhardt, H. (2009 йил 21-Dec). The multi-domain protein Np95 connects DNA methylation and histone modification. *Nucleic Acids Research*.

Rougier, N., Bourc'his, D., Gomes, D., Niveleau, A., Plachot, M., Paldi, A., et al. (1998). Chromosome methylation patterns during mammalian preimplantation development. *Genes & Development*, 12 (14), 2108-13.

Rountree, M., Bachman, K., & Baylin, S. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nature Genetics*, 25 (3), 269-77.

Ruzov, A., Dunican, D., Prokhortchouk, A., Pennings, S., Stancheva, I., Prokhortchouk, E., et al. (2004). Kaiso is a genome-wide repressor of transcription that is essential for amphibian development. *Development (Cambridge, England)*, 131 (24), 6185-94.

Ruzov, A., Shorning, B., Mortusewicz, O., Dunican, D., Leonhardt, H., & Meehan, R. (2009). MBD4 and MLH1 are required for apoptotic induction in xDNMT1-depleted embryos. *Development (Cambridge, England)*, 136 (13), 2277-86.

Sansom, O., Zabkiewicz, J., Bishop, S., Guy, J., Bird, A., & Clarke, A. (2003). MBD4 deficiency reduces the apoptotic response to DNA-damaging agents in the murine small intestine. *Oncogene*, 22 (46), 7130-6.

Santoro, R., & Grummt, I. (2001). Molecular mechanisms mediating methylation-dependent silencing of ribosomal gene transcription. *Molecular Cell*, 8 (3), 719-25.

Santos, F., Hendrich, B., Reik, W., & Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. *Developmental Biology*, 241 (1), 172-82.

Sasai, N., Matsuda, E., Sarashina, E., Ishida, Y., & Kawaichi, M. (2005). Identification of a novel BTB-zinc finger transcriptional repressor, CIBZ, that interacts with CtBP corepressor. *Genes to cells : devoted to molecular & cellular mechanisms*, 10 (9), 871-85.

Sasai, N., Nakao, M., & Defossez, P.-A. (2010). Sequence-specific recognition of methylated DNA by human zinc-finger proteins. *Nucleic Acids Research*, 38 (15), 5015-22.

Sasaki, H., & Matsui, Y. (2008). Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nature reviews Genetics*, 9 (2), 129-40.

Schaefer, M., & Lyko, F. (2010). Lack of evidence for DNA methylation of Invader4 retroelements in Drosophila and implications for Dnmt2-mediated epigenetic regulation. *Nature genetics*, 42 (11), 920-1.

Schaefer, M., Pollex, T., Hanna, K., Tuorto, F., Meusburger, M., Helm, M., et al. (2010). RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes & Development*, 24 (15), 1590-5.

Schmitz, K.-M., Mayer, C., Postepska, A., & Grummt, I. (2010). Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes & Development*, 24 (20), 2264-9.

Screatton, R., Kiessling, S., Sansom, O., Millar, C., Maddison, K., Bird, A., et al. (2003). Fas-associated death domain protein interacts with methyl-CpG binding domain protein 4: a potential link between genome surveillance and apoptosis. *Proceedings of the National Academy of Sciences of the United States of America*, 100 (9), 5211-6.

Segal, E., & Widom, J. (2009). What controls nucleosome positions? *Trends in genetics : TIG*, 25 (8), 335-43.

Sharif, J., Muto, M., Takebayashi, S.-I., Suetake, I., Iwamatsu, A., Endo, T., et al. (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature*, 450 (7171), 908-12.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J., Cole, P., et al. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*, 119 (7), 941-53.

Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T., Webster, Z., et al. (2003). Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Developmental cell*, 4 (4), 481-95.

Skene, P., Illingworth, R., Webb, S., Kerr, A., James, K., Turner, D., et al. (2010). Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. *Molecular Cell*, 37 (4), 457-68.

Smeets, D., Moog, U., Weemaes, C., Vaes-Peters, G., Merkx, G., Niehof, J., et al. (1994). ICF syndrome: a new case and review of the literature. *Human Genetics*, 94 (3), 240-6.

Smith, G. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science (New York, NY)*, 228, 1315-1317.

Song, J., Rechkoblit, O., Bestor, T., & Patel, D. (2011). Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation. *Science (New York, NY)*, 331 (6020), 1036-40.

Soubry, A., Staes, K., Parthoens, E., Noppen, S., Stove, C., Bogaert, P., et al. (2010). The transcriptional repressor Kaiso localizes at the mitotic spindle and is a constituent of the pericentriolar material. *PloS one*, 5 (2), e9203.

Spada, F., Haemmer, A., Kuch, D., Rothbauer, U., Schermelleh, L., Kremmer, E., et al. (2007). DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells. *The Journal of cell biology*, 176 (5), 565-71.

Steenman, M., Rainier, S., Dobry, C., Grundy, P., Horon, I., & Feinberg, A. (1994). Loss of imprinting of IGF2 is linked to reduced expression and abnormal methylation of H19 in Wilms' tumour. *Nature Genetics*, 7 (3), 433-9.

Sterner, D., & Berger, S. (2000). Acetylation of histones and transcription-related factors. *Microbiology and molecular biology reviews : MMBR*, 64 (2), 435-59.

Straussman, R., Nejman, D., Roberts, D., Steinfeld, I., Blum, B., Benvenisty, N., et al. (2009). Developmental programming of CpG island methylation profiles in the human genome. *Nature Structural & Molecular Biology*, 16 (5), 564-71.

Strohner, R., Nemeth, A., Jansa, P., Hofmann-Rohrer, U., Santoro, R., Längst, G., et al. (2001). NoRC--a novel member of mammalian ISWI-containing chromatin remodeling machines. *The EMBO journal*, 20 (17), 4892-900.

Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H., & Tajima, S. (2004). DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. *The Journal of biological chemistry*, 279 (26), 27816-23.

Surani, M., Barton, S., & Norris, M. (1984). Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature*.

Szwagierczak, A., Bultmann, S., Schmidt, C., Spada, F., & Leonhardt, H. (2010). Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Research*, 38 (19), e181.

Tahiliani, M., Koh, K., Shen, Y., Pastor, W., Bandukwala, H., Brudno, Y., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science (New York, NY)*, 324 (5929), 930-5.

Takahashi, Y., Suzuki, C., & Kamataki, T. (1998). Silencing of CYP1A1 expression in rabbits by DNA methylation. *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS*, 247 (2), 383-6.

Takebayashi, S.-i., Tamura, T., Matsuoka, C., & Okano, M. (2007). Major and essential role for the DNA methylation mark in mouse embryogenesis and stable association of DNMT1 with newly replicated regions. *Molecular and Cellular Biology*, 27 (23), 8243-58.

- Thomas, J., & Travers, A. (2001). HMG1 and 2, and related 'architectural' DNA-binding proteins. *Trends in biochemical sciences* , 26 (3), 167-174.
- Thomas, J., & Travers, A. (2001). HMG1 and 2, and related 'architectural' DNA-binding proteins. *Trends in biochemical sciences* , 26 (3), 167-74.
- Trotzier, M.-A., Bronner, C., Bathami, K., Mathieu, E., Abbady, A.-Q., Jeanblanc, M., et al. (2004). Phosphorylation of ICBP90 by protein kinase A enhances topoisomerase IIalpha expression. *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS* , 319 (2), 590-5.
- Tsai, M.-C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J., Lan, F., et al. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. *Science (New York, NY)* , 329 (5992), 689-93.
- Tsai, T., Jiang, Y., Bressler, J., Armstrong, D., & Beaudet, A. (1999). Paternal deletion from Snrpn to Ube3a in the mouse causes hypotonia, growth retardation and partial lethality and provides evidence for a gene contributing to Prader-Willi syndrome. *Human molecular genetics* , 8 (8), 1357-64.
- Tu, S., Teng, Y.-C., Yuan, C., Wu, Y.-T., Chan, M.-Y., Cheng, A.-N., et al. (2008). The ARID domain of the H3K4 demethylase RBP2 binds to a DNA CCGCCC motif. *Nature Structural & Molecular Biology* , 15 (4), 419-21.
- Uemura, T., Kubo, E., Kanari, Y., Ikemura, T., Tatsumi, K., & Muto, M. (2000). Temporal and spatial localization of novel nuclear protein NP95 in mitotic and meiotic cells. *Cell structure and function* , 25 (3), 149-59.
- Unoki, M., Nishidate, T., & Nakamura, Y. (2004). ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain. *Oncogene* , 23 (46), 7601-10.
- Vakoc, C., Mandat, S., Olenchok, B., & Blobel, G. (2005). Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Molecular Cell* , 19 (3), 381-91.
- van Overveld, P., Lemmers, R., Sandkuijl, L., Enthoven, L., Winokur, S., Bakels, F., et al. (2003). Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. *Nature Genetics* , 35 (4), 315-7.
- Velasco, G., Hubé, F., Rollin, J., Neuillet, D., Philippe, C., Bouzinba-Segard, H., et al. (2010). Dnmt3b recruitment through E2F6 transcriptional repressor mediates germ-line gene silencing in murine somatic tissues. *Proceedings of the National Academy of Sciences of the United States of America* , 107 (20), 9281-6.
- Viré, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., et al. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* , 439 (7078), 871-4.
- Vissel, B., & Choo, K. (1989). Mouse major (gamma) satellite DNA is highly conserved and organized into extremely long tandem arrays: implications for recombination between nonhomologous chromosomes. *Genomics* , 5 (3), 407-14.
- Waddington, C. (1942). Canalization of development and the inheritance of acquired characters. *Nature* .
- Walsh, C., Chaillet, J., & Bestor, T. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nature Genetics* , 20 (2), 116-7.
- Wang, J., Hevi, S., Kurash, J., Lei, H., Gay, F., Bajko, J., et al. (2009). The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nature Genetics* , 41 (1), 125-9.
- WATSON, J., & CRICK, F. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* , 171 (4356), 737-738.
- Weber, A., Marquardt, J., Elzi, D., Forster, N., Starke, S., Glaum, A., et al. (2008). Zbtb4 represses transcription of P21CIP1 and controls the cellular response to p53 activation. *The EMBO journal* , 27 (11), 1563-74.

- Weber, M., Hellmann, I., Stadler, M., Ramos, L., Pääbo, S., Rebhan, M., et al. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature Genetics* , 39 (4), 457-66.
- Webster, K., O'Bryan, M., Fletcher, S., Crewther, P., Aapola, U., Craig, J., et al. (2005). Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis. *Proceedings of the National Academy of Sciences of the United States of America* , 102 (11), 4068-73.
- Weintraub, H., Davis, R., Lockshon, D., & Lassar, A. (1990). MyoD binds cooperatively to two sites in a target enhancer sequence: occupancy of two sites is required for activation. *Proceedings of the National Academy of Sciences of the United States of America* , 87 (15), 5623-7.
- Wentworth, B., Donoghue, M., Engert, J., Berglund, E., & Rosenthal, N. (1991). Paired MyoD-binding sites regulate myosin light chain gene expression. *Proceedings of the National Academy of Sciences of the United States of America* , 88 (4), 1242-6.
- Whetstine, J., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., et al. (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* , 125 (3), 467-81.
- Williams, K., Christensen, J., Pedersen, M., Johansen, J., Cloos, P., Rappsilber, J., et al. (2011). TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* , 473 (7347), 343-8.
- Wong, E., Yang, K., Kuraguchi, M., Werling, U., Avdievich, E., Fan, K., et al. (2002). Mbd4 inactivation increases Cright-arrowT transition mutations and promotes gastrointestinal tumor formation. *Proceedings of the National Academy of Sciences of the United States of America* , 99 (23), 14937-42.
- Wu, H., D'Alessio, A., Ito, S., Wang, Z., Cui, K., Zhao, K., et al. (2011). Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes & Development* , 25 (7), 679-84.
- Wu, H., Tao, J., Chen, P., Shahab, A., Ge, W., Hart, R., et al. (2010). Genome-wide analysis reveals methyl-CpG-binding protein 2-dependent regulation of microRNAs in a mouse model of Rett syndrome. *Proceedings of the National Academy of Sciences of the United States of America* , 107 (42), 18161-6.
- Yagi, S., Hirabayashi, K., Sato, S., Li, W., Takahashi, Y., Hirakawa, T., et al. (2008). DNA methylation profile of tissue-dependent and differentially methylated regions (T-DMRs) in mouse promoter regions demonstrating tissue-specific gene expression. *Genome research* , 18 (12), 1969-78.
- Yamada, D., Pérez-Torrado, R., Filion, G., Caly, M., Jammart, B., Devignot, V., et al. (2009). The human protein kinase HIPK2 phosphorylates and downregulates the methyl-binding transcription factor ZBTB4. *Oncogene* , 28 (27), 2535-44.
- Yang, C., van der Woerd, M., Muthurajan, U., Hansen, J., & Luger, K. (2011 йил 29-Jan). Biophysical analysis and small-angle X-ray scattering-derived structures of MeCP2-nucleosome complexes. *Nucleic acids research* .
- Yoder, J., & Bestor, T. (1998). A candidate mammalian DNA methyltransferase related to pmt1p of fission yeast. *Human molecular genetics* , 7 (2), 279-84.
- Yoon, H.-G., Chan, D., Reynolds, A., Qin, J., & Wong, J. (2003). N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso. *Molecular Cell* , 12 (3), 723-34.
- Young, J., Hong, E., Castle, J., Crespo-Barreto, J., Bowman, A., Rose, M., et al. (2005). Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. *Proceedings of the National Academy of Sciences of the United States of America* , 102 (49), 17551-8.
- Zemach, A., McDaniel, I., Silva, P., & Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science (New York, NY)* , 328 (5980), 916-9.



- Zhang, C., McKinsey, T., & Olson, E. (2002). Association of class II histone deacetylases with heterochromatin protein 1: potential role for histone methylation in control of muscle differentiation. *Molecular and Cellular Biology* , 22 (20), 7302-7312.
- Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S.-L., Chen, H., et al. (2006). Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell* , 126 (6), 1189-201.
- Zhang, Y., & Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes & Development* , 15 (18), 2343-60.
- Zhang, Y., Ng, H., Erdjument-Bromage, H., Tempst, P., Bird, A., & Reinberg, D. (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes & development* , 13 (15), 1924-35.
- Zhang, Y., Rohde, C., Reinhardt, R., Voelcker-Rehage, C., & Jeltsch, A. (2009). Non-imprinted allele-specific DNA methylation on human autosomes. *Genome Biology* , 10 (12), R138.
- Zhao, X., Ueba, T., Christie, B., Barkho, B., McConnell, M., Nakashima, K., et al. (2003). Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function. *Proceedings of the National Academy of Sciences of the United States of America* , 100 (11), 6777-82.
- Zilberman, D., Coleman-Derr, D., Ballinger, T., & Henikoff, S. (2008). Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* , 456 (7218), 125-9.

